基于图像特征的数字水印算法研究

王向阳[1),2)] 邹 俊[1)] 侯丽敏[1)]

[1)](辽宁师范大学计算机与信息技术学院,大连 116029) [2)](苏州大学江苏省计算机信息处理技术重点实验室,苏州 215006)

摘 要 为了使数字水印具有更鲁棒的抗攻击能力,提出了一种基于图像特征的数字水印新算法。该算法首先利用 Harris-Laplace 算子提取载体图像特征点;然后结合特征尺度自适应确定局部特征区域;最后,采纳 DFT 中频幅值比较策略将数字水印信息重复嵌入到多个不相交的局部特征区域中。检测时,根据模糊模式识别的最大隶属度原则检测水印信息。仿真实验结果表明,该新算法不仅具有较好的透明性,而且对常规信号处理(中值滤波、边缘锐化、叠加噪声和 JPEG 压缩等)和去同步攻击(旋转、平移、缩放、行列去除、剪切和局部随机弯曲等)均具有较好的鲁棒性。

关键词 图像水印 去同步攻击 特征点 局部特征区域

中图法分类号: TP309 **文献标识码**: A **文章编号**: 1006-8961(2006)11-1562-04

A New Featured-based Image Watermarking Algorithm

WANG Xiang-yang[1),2)], WU Jun[1)], HOU Li-min[1)]

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

[2)](*Jiangsu Province Key Lab. for Computer Information Processing Technology, Soochow University, Suzhou* 215006)

Abstract In this paper, a novel feature-based image watermarking scheme is proposed. Firstly, the Harris-Laplace detector is utilized to extracted feature points, which can survive a variety of attacks. Then, nonoverlapped disks around local feature points are defined adaptively according to the feature scales. Finally, several copies of the digital watermark are embedded into the nonoverlapped disks by comparing the DFT mid-frequency magnitudes. In watermark detection, the digital watermark can be extracted by using the maximum membership criterion. Experimental results show that the proposed scheme is not only invisible and robust against common signals processing such as median filtering, sharpening, noise adding, and JPEG compression, but also robust against the desynchronization attacks such as rotation, translation, scaling, row or column removal, shearing, and local random bend.

Keywords image watermarking, desynchronization attack, feature point, local feature region

1 引 言

数字图像水印技术发展到今天,已有大量不同的算法,虽然它们都广泛提出了“鲁棒性”声明,但遗憾的是,现有绝大多数图像水印算法仅仅能够对抗常规的信号处理(如压缩、滤波、噪声干扰等)和简单的全局仿射变换(包括旋转、缩放和平移,即都属于最基本的去同步攻击),而无法有效抵抗一般性去同步攻击,如剪切(shearing)、行列去除(column or line removal)、局部弯曲(random bend)等等[1,2]。也就是说,现有数字图像水印技术抵抗去同步攻击的能力都很差。因此,抗去同步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。所谓去同步攻击(desynchronization attack),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(即改变水印的嵌入位置),以导致检测器找不到有效水印[3,4]。去同步攻击包括全局仿射变换、局部随机弯曲、剪切、行列去除、几何变换组合等多种形式。

基金项目:辽宁省自然科学基金项目(20032100);视觉与听觉信息处理国家重点实验室(北京大学)开放基金项目(0503)

收稿日期:2006-05-08;**改回日期**:2006-08-03

第一作者简介:王向阳(1965~),男,教授。于 1995 年获吉林大学工学硕士学位。主要研究领域为多媒体信息处理技术、网络信息安全技术。E-mail: wxy37@126.com

本文以 Harris-Laplace 算子为基础,提出了一种基于图像特征点的可有效抵抗去同步攻击的图像水印新算法。该算法首先利用 Harris-Laplace 算子提取载体图像的特征点;然后结合特征尺度自适应确定局部特征区域;最后采纳 DFT(discrete Fourier transform)中频幅值比较策略来将水印信息重复嵌入到多个不相交的局部特征区域内。

2　数字水印的嵌入

由数字图像水印系统的通信模型(将载体图像看成信道,将数字水印看作被传输信息,而将各种有意或无意攻击当作噪声干扰)知:当原始载体图像被划分成若干局部特征区域以后,则所有局部特征区域均可当作是物理上分布而逻辑上统一的传输信道群。因此,在数字水印信息传输过程中,即使部分信道遭到破坏,其余信道仍可以保证水印信息的正常传输。基于上述思想,本文将采用冗余嵌入策略,以提高整个水印系统的鲁棒性,即将同一数字水印重复地嵌入到所有的局部特征区域内。整个数字水印嵌入方案(关键步骤)可描述如下。

(1)由密钥 Key1 产生一个大小为 L 的双极性序列 $\boldsymbol{W}=\{w_i, i=1,\cdots,L\}$,并将其作为数字水印。其中,$w_i\in\{-1,1\}$ 且满足 $\sum_{i=1}^{L} w_i=0$(均值为0),以使含水印图像的整体能量保持均衡,即透明性较好。

(2)利用 Harris-Laplace 算子从原始载体中提取图像特征点[5],以得到图像特征点集 $\boldsymbol{P}=\{p_i, i=1,\cdots,n\}$。

(3)以图像特征点为中心,利用基于特征尺度的局部特征区域确定方法对载体图像进行分割[6],以得到一系列局部特征区域(即圆片)$\boldsymbol{O}=\{o_k, k=1,\cdots,m\}$。

(4)从局部特征区域集 $\boldsymbol{O}$ 中取出一局部特征区域 o_k,并将其四周"补0"以得到方形子图像;再通过对方形子图像做中心化 DFT 变换(变换原点为子图像中心)来得到第 k 个局部特征区域中心化频谱 F_k,并取出其幅值谱 M_k 和相位谱 φ_k。

(5)采用 DFT 幅值比较策略在局部特征区域 o_k 中嵌入数字水印,具体操作过程为:

首先,在幅值谱 M_k 内选择半径 r_1 和 r_L,并满足 $r_1<r_L$,以便使 r_1 和 r_L 之间的环形区域覆盖中频带。设 $\{\boldsymbol{C}(r_i), i=1,\cdots,L\}$ 是中频带内半径由小到大的同心圆族,且满足 $r_1\leqslant r_i\leqslant r_L$。设 $\boldsymbol{C}(r_i)$ 上有 m_i 个成对出现的幅值谱系数(它们与中心成 90°),若将它们分成 $m_i/2$ 组,则其可表示为

$$\{(M_k(x_i^{(s)},y_i^{(s)}),M_k(y_i^{(s)},-x_i^{(s)})), s=1,\cdots,m_i/2\} \tag{1}$$

然后,通过密钥 Key2 在环形区域内选择 L 个起始点,表示为 $\{(x_i^{(1)},y_i^{(1)}), i=1,\cdots,L\}$,并使这些点均匀分布于同心圆族之上。

最后,从起始点开始,通过修改 $M_k(x_i^{(s)},y_i^{(s)})$ 与 $M_k(y_i^{(s)},-x_i^{(s)})$ 间的相对大小来嵌入水印信息位,即将 1bit 水印信息 w_i 嵌入到 $\boldsymbol{C}(r_i)$ 中。水印嵌入方法如下:

如果 $w_i=1$,则按下面规则对 $\boldsymbol{C}(r_i)$ 上所有成对出现的幅值谱系数按顺时针顺序依次进行修改

$$\begin{cases} M_k(x_i^{(s)},y_i^{(s)})-M_k(y_i^{(s)},-x_i^{(s)})\geqslant\alpha \\ M_k(x_i^{(s)},y_i^{(s)})-M_k(y_i^{(s)},-x_i^{(s)})<\alpha \end{cases} \tag{2}$$

如果 $w_i=-1$,则按下面规则修改幅值谱系数

$$\begin{cases} M_k(x_i^{(s)},y_i^{(s)})-M_k(y_i^{(s)},-x_i^{(s)})\leqslant-\alpha \\ M_k(x_i^{(s)},y_i^{(s)})-M_k(y_i^{(s)},-x_i^{(s)})>-\alpha \end{cases} \tag{3}$$

其中,α 代表水印嵌入强度。

重复上述过程,直到将 L bits 水印信息全部嵌入到环形区域内为止,并最终得到嵌入水印的幅值谱 $\hat{M}_k$。

(6)由于上述嵌入方法并未改变 DFT 频谱的对称性,因此可以先直接将含水印幅值谱 $\hat{M}_k$ 及原始相位谱 φ_k 用来组成新的 DFT 频谱系数 $\hat{F}_k$,然后对 $\hat{F}_k$ 进行 IDFT 变换,便可以得到含有水印的方形子图像,再将其四周"去0"后,即得到含水印的局部特征区域 $\hat{o}_k$。

(7)重复步骤(4)~步骤(6),直到所有的局部特征区域都被嵌入水印为止。

3　数字水印的检测

由于水印信息被重复地嵌入到不同的信道(局

部特征区域)中,而且信道间彼此独立。因此,可采用相同的方法将待检测图像划分成若干个局部特征区域,且只要两个以上圆片能检测到水印,便可认为水印存在于待检测图像中。整个数字水印的检测过程如下:

(1)使用与嵌入过程相同的密钥 Key1 来产生原始水印序列 $\boldsymbol{W}=\{w_i,i=1,\cdots,L\}$。

(2)利用 Harris-Laplace 算子从待检测图像中提取图像特征点[5],以得到图像特征点集 $\tilde{\boldsymbol{P}}=\{\tilde{p}_i,i=1,\cdots,n\}$。

(3)以图像特征点为中心,利用基于特征尺度的局部特征区域确定方法[6]对待检测图像进行分割,以得到一系列局部特征区域(即圆片)$\tilde{\boldsymbol{O}}=\{\tilde{o}_k,k=1,\cdots,m\}$。一般说来,含水印图像遭受攻击以后,检测所得到的新局部特征区域可能会发生变化,但至少有一部分含水印的局部特征区域仍存在于 $\tilde{\boldsymbol{O}}$ 内。

(4)从局部特征区域集 $\tilde{O}$ 中取出一局部特征区域 $\tilde{o}_k$,先将其四周"补 0"以得到方形子图像;再通过对方形子图像做中心化 DFT 变换,以得到中心化频谱 $\tilde{F}_k$,并取出其幅值谱 $\tilde{M}_k$ 和相位谱 $\tilde{\varphi}_k$。

(5)利用幅值谱 $\tilde{M}_k$ 旋转的角度来对局部特征区域进行角度修正。由于频域第 10 行中的最大数值基本能代表幅值谱中"十字线"的位置,故可利用它来计算偏角 θ,并将局部特征区域逆时针旋转 $\theta°$。局部特征区域四周"补 0"后是一个尺寸为 $2R\times 2R$ 的方形子图像,设频域第 10 行中最大数值位于第 h 列,则可按如下规则计算偏角 θ。

$$\theta=\begin{cases}90°-\arctan((R-h)/(R-10))\times 180°/\pi, & h<R\\ \arctan((R-h)/(R-10))\times 180°/\pi, & h\geqslant R\end{cases} \tag{4}$$

(6)由 DFT 性质知,图像经历几何变换后,其频域上的嵌入区域大小保持不变。故可以先使用与嵌入时相同的密钥 Key2 在环形区域 $[r_1,r_L]$ 内确定 L 个起始位置。然后从起始点开始,沿延顺时针顺序依次对 $\boldsymbol{C}(r_i)$ 中所有成对出现的幅值谱系数按如下规则提取水印信息:

$$\tilde{w}_i^{(s)}=\begin{cases}1, & \tilde{M}_k(x_i^{(s)},y_i^{(s)})-\tilde{M}_k(y_i^{(s)},-x_i^{(s)})\geqslant 0\\ -1, & \tilde{M}_k(x_i^{(s)},y_i^{(s)})-\tilde{M}_k(y_i^{(s)},-x_i^{(s)})<0\end{cases} \tag{5}$$

其中,$\tilde{w}_i^{(s)}(s=1,\cdots,m_i/2)$ 表示从 $\boldsymbol{C}(r_i)$ 中提取的水印系数(共 $m_i/2$ 组)。

一般说来,待检测图像遭受攻击以后,其幅值谱可能会发生微小变化,为此本文将采用模糊模式识别的最大隶属度原则来恢复数字水印的位信息。

根据隶属度相关理论,本文将 $\tilde{w}_i^{(s)}$ 对 1 的隶属度定义为

$$d(1)=\frac{2\times N_1}{m_i}$$

其中,N_1 是 $\boldsymbol{W}$ 中 $\tilde{w}_i^{(s)}$ 值为 1 的个数。

而 $\tilde{w}_i^{(s)}$ 对 -1 的隶属度定义为

$$d(-1)=\frac{2\times N_2}{m_i}$$

其中,N_2 是 $\boldsymbol{W}$ 中 $\tilde{w}_i^{(s)}$ 值为 -1 的个数。

根据模糊模式识别的最大隶属度原则,就可以通过比较 $d(1)$ 和 $d(-1)$ 来决策 $\boldsymbol{C}(r_i)$ 上的最终提取结果:

$$\tilde{w}_i=\begin{cases}1, & d(1)\geqslant d(-1)\\ -1, & d(1)<d(-1)\end{cases} \tag{6}$$

其中,$\tilde{w}_i$ 表示从 $\boldsymbol{C}(r_i)$ 中提取的 1bit 水印信息。重复上述过程,直到整个环形区域被处理完,即可得到所提取出的数字水印 $\tilde{\boldsymbol{W}}=\{\tilde{w}_i,i=1,\cdots,L\}$。显然,尽管图像遭受各种攻击后,其幅值谱可能发生变化,但利用最大隶属度原则,仍可以较为准确地提取出数字水印。

(7)重复步骤(4)~步骤(6),直到所有局部特征区域检测完毕为止。只要有两个以上的局部特征区域成功检测到数字水印,便可认为待检测图像中存在数字水印,即检测成功;否则,检测失败。

4 仿真实验与结论

为了验证本文算法的高效性,以下给出了性能检测与抗攻击能力的测试结果。实验中,所选用的原始载体分别为 512 × 512 × 8bit 标准灰度图像 Lena、Mandrill 和 Pepper,数字水印采用了 32bit 的二元随机序列。另外,自适应常数选取为 $\tau=9$,水印嵌入强度选取为 $\alpha=20$。

表 1 和表 2 给出了本文算法的鲁棒性能测试结果(其中,分子表示从攻击后的含水印图像中成功检测到数字水印的局部特征区域数目,而分母则表示载体图像中嵌有水印的局部特征区域数目)。

本文提出了一种基于内容特征的可有效抵抗去同步攻击的数字图像水印方案,其主要包括如下特点:(1)所提取的图像特征点不仅稳定性好,而且分布均匀,还提高了整个水印系统对常规信号处理及

表 1　对常规处理的抵抗能力(重构率)

Tab. 1　Fraction of correctly detected watermark under common signal processing(detection rates)

攻击方式		图像		
		Lena	Mandrill	Pepper
无攻击		6/6	11/12	8/8
中值滤波(3×3)		3/6	7/12	4/8
锐化(3×3)		3/6	6/12	5/8
叠加高斯噪声		2/6	4/12	4/8
JPEG 压缩	70	4/6	9/12	7/8
	50	4/6	8/12	6/8
	30	2/6	8/12	4/8
中值滤波 + JPEG90		3/6	7/12	4/8
锐化(3×3) + JPEG90		3/6	6/12	6/8

表 2　对去同步攻击的抵抗能力(重构率)

Tab. 2　Fraction of correctly detected watermark under desynchronization attacks(detection rates)

攻击方式		Lena	Mandrill	Pepper
去除 8 行 16 列		5/6	7/12	6/8
中心裁掉 10%		3/6	5/12	4/8
剪切成原先的 55%		4/6	6/12	6/8
旋转 5°		4/6	5/12	4/8
平移(水平 10,垂直 10)		5/6	10/12	5/8
缩放	0.8	1/6	2/12	3/8
	0.9	3/6	5/12	3/8
	1.2	1/6	1/12	2/8
局部弯曲		3/6	7/12	6/8
中心裁掉 10% + JPEG70		2/6	5/12	4/8
旋转 5° + 缩放 0.9		3/6	5/12	5/8
平移(水平 10,垂直 10) + 旋转 5° + 缩放 0.9		2/6	3/12	1/8

去同步攻击的抵抗能力;(2)能够结合图像内容自适应确定局部特征区域尺寸,从而大大增强了图像水印系统的工作性能;(3)结合最大隶属度理论恢复数字水印信息,可以降低了水印系统的虚警率。

参考文献(References)

1　Barni M, Cox I J, Kalker T. Digital watermarking [A]. In: Proceedings of 4th International Workshop on Digital Watermarking (IWDW), Siena, Italy, 2005: 15 ~ 19.

2　Li Chang-li, Lu Zhao-yang. Desynchronization attack on digital watermarks and their countermeasures [J]. Journal of Image and Graphics, 2005, **10**(4): 403 ~ 409. [李昌利, 卢朝阳. 数字水印的去同步攻击及其对策[J]. 中国图象图形学报, 2005, **10**(4): 403 ~ 409.]

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

4　Liu Jiu-fen, Huang Da-ren, Huang Ji-wu. Survey on watermarking against geometric attack [J]. Journal of Electronics and Information Technology, 2004, **26**(9): 1495 ~ 1503. [刘九芬, 黄达人, 黄继武. 图像水印抗几何攻击研究综述[J]. 电子与信息学报, 2004, **26**(9): 1495 ~ 1503.]

5　Mikolajczyk K, Schmid C. Scale & Affine Invariant Interest Point Detectors [J]. International Journal of Computer Vision, 2004, **60**(1): 63 ~ 86.

6　Manjunath B S, Shekhar C, *et al*. A new approach to image feature detection with applications[J]. Pattern Recognition, 1996, **29**(4): 627 ~ 640.