第 26 卷第 12 期 电子与信息学报 Vol.26No.12
2004 年 12 月 Journal of Electronics & Information Technology Dec. 2004

抵抗几何攻击的数字图像水印 [1]

桑茂栋 赵 耀

(北京交通大学信息科学研究所 北京 100044)

摘 要: 该文实现了一种抵抗几何攻击的数字图像水印方案. 水印信号为根据密钥产生的服从正态分布的实数序列, 对水印信号作预处理并进行周期化之后在空域将其嵌入数字图像中. 检测端不需要原始图像, 根据预测水印的自相关函数 (ACF) 判断水印图像遭受的几何变换攻击并进行逆变换, 然后进行水印检测. 实验表明该方案对压缩、滤波、剪切等常见的图像处理攻击同样具有很高的鲁棒性.

关键词: 数字图像水印, 版权保护, 鲁棒性, HVS, 几何攻击, 仿射变换

中图分类号: TP391 **文献标识码**: A **文章编号**: 1009-5896(2004)12-1875-07

Digital Image Watermarking Resisting to Geometrical Attacks

Sang Mao-dong Zhao Yao

(*Institute of Info. Science, Beijing Jiaotong University, Beijing* 100044, *China*)

Abstract In this paper a watermarking scheme for digital image that can resist geometrical attacks is proposed. The watermark is a key-dependent random vector according to Gaussian distribution. After preprocessed and periodized, the watermark is embedded into cover image in spatial domain exploiting the properties of HVS. The cover image is not needed in watermark detection process. First, compute the AutoCorrelation Function(ACF) of estimated watermark. The peaks of ACF are used to determine the geometrical attacks applied to the stegoimage. Experimental results show the scheme is robust to compression, filtering, cropping as well as geometrical attacks.

Key words Digital image watermarking, Copyright protection, Robustness, HVS, Geometrical attacks, Affine transformation

1 引言

作为传统加密方法的有效补充手段, 数字水印 (Digital watermarking) 技术被认为是解决数字化时代数字作品版权保护的一个重要手段, 并成为多媒体信号处理领域的一个研究热点. 数字水印技术是指用信号处理的方法在数字化的多媒体数据中嵌入隐蔽的标记, 这种标记通常是不可见的, 只有通过专用的检测器或阅读器才能提取. 衡量数字水印技术先进可靠的标准是能否抵抗各种水印攻击. 目前已经提出的许多水印处理算法 (详细描述见文献 [1]) 只能抵抗几种常见的攻击, 如对图像进行格式变换、低通滤波、数据压缩等, 而对于诸如放缩、剪切、旋转、拉伸、长宽比改变等几何攻击则无能为力, 所以几何攻击被认为是数字水印技术走上商用的瓶颈.

当前针对抵抗几何攻击的水印方法主要有 3 种: (1) 借助原始图像 [2,3]; (2) 在 RST 变换的不变域进行水印嵌入 [4,5]; (3) 将几何攻击模拟为仿射变换, 在检测过程中利用嵌入的特定模板预测遭受的仿射变换 [6]. 此外, Kutter 在文献 [7] 中提出了一种自参考方法 (Self-reference), 将

[1] 2003-07-05 收到, 2004-01-22 改回

国家自然科学基金 (60172062, 60373028)、霍英东青年教师基金 (81053) 和留学回国人员科研启动基金资助课题

同一个水印在图像 4 个不同的位置分别嵌入，检测的时候计算预测水印的自相关函数 (ACF)，根据 ACF 峰值位置的改变预测水印图像遭受的仿射变换的参数。这种水印算法可以较好地抵抗旋转、伸缩、平移和改变长宽比等全局几何攻击，但是由于峰值太少且幅度下降明显，不能很好地抵抗剪切、压缩、滤波和镜像等攻击。

本文主要考虑抗几何攻击的盲检测图像水印算法，我们利用 Kutter 自参考水印的思想以提高鲁棒性为着眼点提出了改进算法，我们采用服从高斯分布的实数序列作为水印信号，不是将水印进行多次嵌入而是将水印进行周期化处理，为了更好地跟踪图像经历的攻击，我们将水印扩展到整幅图像大小，这样处理后的水印自相关函数将会产生大量峰值，而且这些峰值位置构成形状规则的栅格。检测的时候不需要原始图像，先做水印预测然后计算预测水印的自相关函数，通过自相关函数的峰值位置构成的形状规则的栅格与原始栅格的比较，判断遭受的几何变换攻击并进行逆变换，最后根据常规的相关检测方法进行水印检测。本文第 2 节介绍水印的产生、构造及其基于人眼视觉系统 (HVS) 特性的自适应嵌入。第 3 节介绍如何进行水印预测和判断几何变换攻击以及水印检测。第 4 节对该方法做了鲁棒性实验并给出实验结果。第 5 节总结全文。

2 水印的嵌入过程

水印的嵌入框图如图 1 所示。

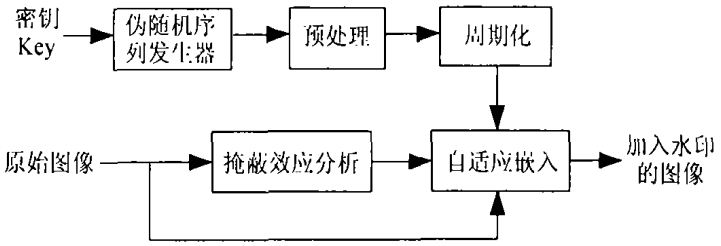

图 1 水印的嵌入框图

2.1 水印的产生和构造

为了增加水印算法的安全性以及抵抗共谋攻击的鲁棒性，我们采用服从高斯分布 $N(0,\sigma_n^2)$ 的实数序列 $\{p_1,p_2,\cdots,p_n\}$ 作为水印。我们对水印序列作如下预处理：将上述序列扩展为二维矩形块 $N_1\times N_2(n=N_1\times N_2)$，然后借鉴文献 [8] 的思想，在水平、垂直、对角线三个方向上分别翻转复制一次，形成中心对称的宏块 $2N_1\times 2N_2$(如图 2 所示)，这样处理可以增加抵抗镜像攻击的鲁棒性。为了更好地跟踪图像经历的攻击，我们将产生的宏块进行多次复制，使之形成与原始图像大小相同的周期性水印 w(行周期 $T_1=2N_1$，列周期 $T_2=2N_2$)。水印产生和构造框图如图 3 所示。

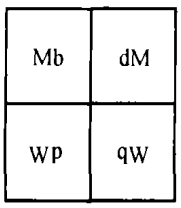

图 2 宏块的形状

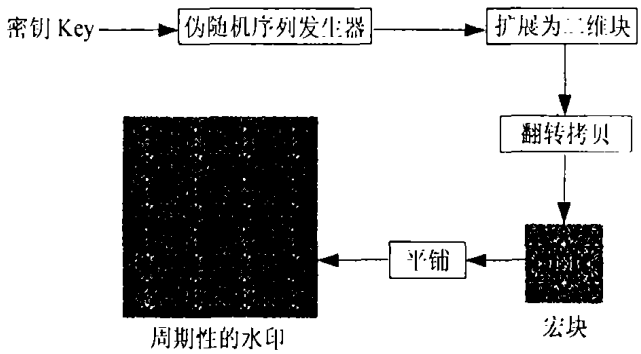

图 3 水印的预处理及周期化

2.2 水印的嵌入

为了在保证水印不可见性的前提下增加水印的强度，我们考虑 HVS 特性，在图像的纹理区域和平滑区域以不同强度嵌入水印。文献 [9] 提出了一种纹理掩蔽效应函数称为噪音可见函数 NVF(Noise Visibility Function) 。在此，我们使用广泛应用于图像重建 [10] 中著名的 NVF 形式：

$$\mathrm{NVF}(i,j)=\frac{1}{1+\theta\sigma_x^2(i,j)} \tag{1}$$

其中 θ 是调整参数 (Tuning parameter)，每幅图像有不同的调整参数，将其写为如下形式：

$$\theta=D/\sigma_{x\max}^2 \tag{2}$$

其中 $\sigma_{x\max}^2$ 是给定图像局部方差的最大值，$D\in[50,1000]$ 是经验值。实验中我们取 150 。

基于以上 HVS 模型，我们可以写出嵌入等式：

$$y(i,j)=x(i,j)+(1-\mathrm{NVF}(i,j))\cdot w(i,j)\cdot\alpha \tag{3}$$

其中 α 表示水印强度。由于非常平滑的区域的 NVF 接近 1，水印的嵌入强度则趋于 0，导致嵌入平滑区域的水印信息很少，这对后面的水印估计会产生不利影响。为了避免这个问题，我们做了一下改变，在人眼可见阈值的范围内增加平滑区域的水印强度：

$$y(i,j)=x(i,j)+(1-\mathrm{NVF}(i,j))\cdot w(i,j)\cdot\alpha+\mathrm{NVF}(i,j)\cdot w(i,j)\cdot\beta \tag{4}$$

其中，α,β 表示水印强度，一般 β 取值小于 α 。

3 水印估计和检测

水印检测框图如图 4 所示。

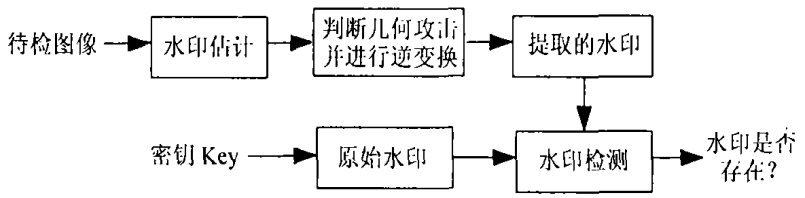

图 4　水印的检测框图

3.1 水印估计

假设图像和水印独立同分布，即 $x\sim N(\overline{x},\boldsymbol{R}_x)$，$w\sim N(0,\boldsymbol{R}_w)$，此时，水印的方差矩阵 $\boldsymbol{R}_w$ 包含掩蔽效应模型调制的影响，我们可以确定 [9]：

$$\hat{w}=\frac{\boldsymbol{R}_w}{\boldsymbol{R}_w+\boldsymbol{R}_x}(y'-\overline{y}') \tag{5}$$

其中假设 $\overline{y}'\approx\overline{x}$，$\overline{y}'$ 表示受到攻击的水印图像的局部均值，$\hat{\boldsymbol{R}}_x=\max(0,\hat{\boldsymbol{R}}_y-\boldsymbol{R}_w)$ 是图像局部方差 ($\boldsymbol{R}_x=\sigma_x^2\boldsymbol{I}$) 的 ML 估计。

实验中，为了减少水印估计时水印图像的边缘区域对估计水印的影响，我们使用边缘保护特性比较好的中值预测 (大小 3×3 的窗口) 估计水印图像的局部均值 $\overline{y}'$，用水印图像 y' 减去局部均值图像 $\overline{y}'$ 所得高频部分的 NVF 代替 $\boldsymbol{R}_w/(\boldsymbol{R}_w+\boldsymbol{R}_x)$，取得良好效果，估计的水印为

$$\hat{w}=\mathrm{NVF}(y'-\overline{y}')\cdot(y'-\overline{y}') \tag{6}$$

3.2 判断几何攻击

常用的几何攻击例如旋转、放缩、平移、长宽比改变、拉伸、扭曲等都可以用仿射变换 A 表示. 设坐标为 (i,j) 的点经过仿射变换 A 之后变为 (i',j')，则

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = A \cdot \begin{bmatrix} i \\ j \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}$$

其中 a, b, c, d 表示线性变化矩阵的参数，e, f 表示平移参数. 平移参数可以暂时不用考虑，因为通过计算原始水印和估计水印的互相关可以消除平移对水印检测的影响 (见 3.3 节). 我们主要考虑前 4 个参数的影响.

为了确定这些参数我们需要计算预测水印 $\hat{w}$ 的 ACF:

$$R_{\hat{w},\hat{w}}(m,n) = \sum_i \sum_j \hat{w}(i,j)\hat{w}(i+m,j+n) \tag{7}$$

由于我们构造的水印行和列均具有周期性，其自相关函数则会产生多个峰值并且这些峰值的坐标点组成形状规则的矩形栅格，栅格的行距和列距分别等于水印的行周期 T_1 和列周期 T_2.

当水印经历仿射变换攻击时，栅格的行距、列距以及两者的角度会根据仿射变换的参数做线性变化，但是其规则的周期性结构仍然保持不变，如图 5 所示. 因为水印是扩展到整幅图像的，水印遭受的几何攻击可以完全反映水印图像遭受的几何攻击，因此，我们可以利用水印自相关函数峰值规则的栅格结构，根据栅格行距、列距以及两者角度的变化判断水印图像遭受的几何攻击.

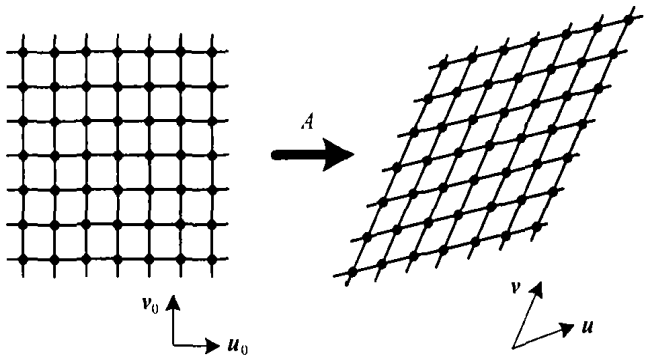

图 5 原始水印及其经仿射变换之后 ACF 峰值组成的栅格

3.3 水印检测

水印检测仍然采用常用的相关检测. 如果水印图像经历了几何变换攻击，首先根据上一节确定的仿射变换参数，对估计的水印 $\hat{w}$ 进行反变换得到最终的预测水印 w^*，然后计算原始水印 w 与水印 w^* 的相似性:

$$\mathrm{sim}(w,w^*) = \frac{\max_{m,n}(R_{w,w^*}(m,n))}{\sqrt{(w^* \cdot w^*)}} \tag{8}$$

其中 $R_{w,w^*}(m,n) = \sum_i \sum_j w(i,j)w^*(i+m,j+n)$，表示原始水印 w 与水印 w^* 的互相关函数，取其最大值是为了消除平移攻击的影响. 如果水印图像没有经历几何攻击，直接对估计水印 $\hat{w}$ 做水印检测.

计算出相似性 $\mathrm{sim}(w,w^*)$ 之后，将其与事先定义的阈值 T 比较，如果 $\mathrm{sim}(w,w^*) > T$，则说明水印存在，否则水印不存在.

4 实验结果

在实验中我们用的图像是大小为 256×256 的灰度图像 Rose，选择的伪随机序列服从标准正态分布，长度为 n=1024。为了方便起见，我们将其扩展为二维方块 ($N_1 = N_2 = 32$)，最后生成的周期性水印 w 的周期为 $T_1 = T_2 = 64$。水印嵌入式 (4) 两个参数选择为 $\alpha = 8, \beta = 3$。嵌入前后的图像如图 6 所示。

图 6　原始图像 Rose(左) 及嵌入水印后的图像 (右) PSNR=33.5009

由于上述周期水印的周期为 $T_1 = T_2 = 64$，所以对于大小为 256×256 的水印来说，其 ACF 峰值应该有 49 个并且 ACF 峰值间隔与水印周期相同。图 7 给出了原始水印的 ACF 峰值及其经过旋转 20° 之后的 ACF 峰值。

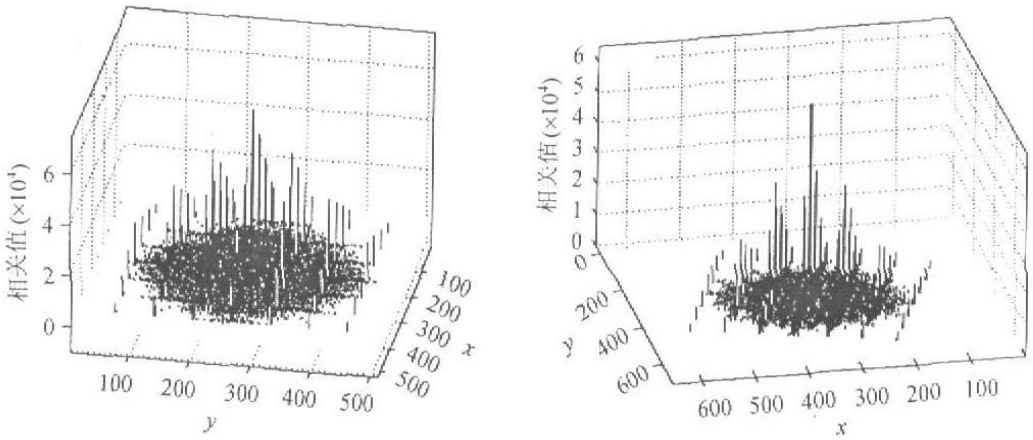

图 7　原始水印的 ACF 峰值 (左) 原始水印经过旋转 20° 的 ACF 峰值 (右)

做水印检测的时候，为了节省计算时间，我们采取如下措施：计算原始宏块 (64×64) 与提取出的水印 w^* 的互相关函数，由于水印周期为 64，所以互相关函数会产生 16 个峰值，我们取其最大值乘以 16 作为式 (8) 的分子，分母保持不变。

我们对嵌入水印的图像做了安全性检测实验。用 1000 个不同的随机水印分别对其进行测试，结果使用原始水印的输出结果 (152.5193) 远远大于使用其他水印的检测结果 (均值为 18.2881，最大为 24.4989)。根据 Cox[2] 的分析，随机序列长度为 1000 时，选取的阈值至少应该比其他水印检测输出结果的平均值大 6。因此，我们可以选择阈值 T=25。

我们对嵌入后的水印图像分别做旋转、放缩、拉伸、扭曲、行列删除、剪切等几何攻击和滤波攻击。对水印图像进行这些攻击之后，水印检测的结果如表 1 所示。最后，我们对水印图像进行 JPEG 压缩攻击，图 8 给出了在不同压缩质量下的检测结果。

表 1 对水印图像进行几何攻击和滤波攻击后的水印检测结果

水印图像遭受的攻击	水印检测的结果	水印图像遭受的攻击 (滤波)	水印检测的结果
旋转 (顺时针 20°)	107.7105	高斯低通 (3×3 窗口, 标准差 0.5)	131.3307
放大 (320×320)	152.4250		
缩小 (128×128)	71.8195		
行 / 列删除 (64, 128, 192 三行三列)	136.7596	中值滤波 (窗口大小 3×3)	29.6647
剪切 (图 9)	116.2540/ 94.1796	维纳滤波 (窗口大小 3×3)	45.6594
翻转 (水平和垂直均翻转)	152.4250	均值滤波 (窗口大小 3×3)	36.5210
拉伸 / 扭曲 (垂直拉伸 150%, 水平扭曲 20°)	94.3480	Motion 滤波 (水平 9 像素)	46.1210

需要说明的是：在对遭受缩小攻击的水印图像进行检测的时候，我们没有将其恢复为原始大小，而是对原始水印做了同样的缩小变换，然后进行水印检测。在对遭受剪切攻击的水印图像的时候按照式 (8) 进行水印检测，而没有采取针对其他攻击的检测措施。

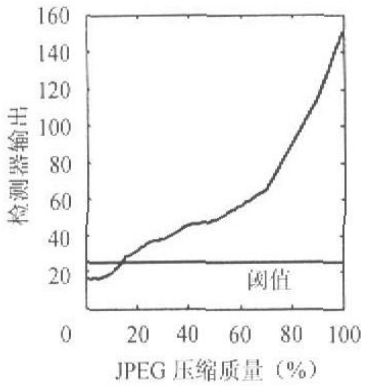

图 8 JPEG 压缩的检测结果

图 9 剪切攻击后的水印图像
(左 256×256, 右 180×180)

5 结论

本文在 Kutter 提出的对同一个水印进行多次嵌入的基础上，引入了周期水印的方法，提出了一种能够抵抗几何攻击的自适应图像水印方案。水印信号为服从标准正态分布的实数序列，我们先将其设计成一种中心对称的二维周期水印，进行周期化之后利用 HVS 特性在空间域将周期水印嵌入到原始图像中，在算法的复杂性、水印的鲁棒性 (嵌入强度) 和不可见性之间达到了一种平衡。实验结果表明该方案对常见的图像处理操作也具有很好的鲁棒性，即使水印图像遭受针对水印估计的中值滤波攻击之后，仍然能够检测到水印的存在。本方案下一步的方向应该是：(1) 根据原始图像的随机模型进一步挖掘图像的局部特征，利用图像的局部特征，结合人类视觉模型的特性，提高掩蔽效应模型的性能，在保证水印不可见性的前提下进一步提高水印的鲁棒性；(2) 根据嵌入模型以及图像、水印的统计模型，采取相应的措施提高水印估计和提取的性能从而提高水印的检测性能。

参 考 文 献

[1] Swanson M D, Kobayashi M, Tewfik A H. Multimedia data embedding and watermarking techniques. *Proc. IEEE*, 1998, 86(6): 1064–1087.

[2] Cox I J, Kilian J, Leighton T, Shamoon T. Secure spread spectrum watermarking for multimedia. *IEEE Trans. on Image Processing*, 1997, 6(12): 1673–1687.

[3] Bas P, Chassery J-M, Macq B. Geometrically invariant watermarking using feature point. *IEEE Trans. on Image Processing*, 2002, 11(9): 1014–1028.

[4] Oruanaidh J, Pun T. Rotation, scale and translation invariant spread spectrum digital image watermarking. *Signal Processing*, 1998, 66(3): 303–317.

[5] Lin C Y, Bloom J A, Cox I J, Miller M L, Lui Y M. Rotation, scale, and translation resilient watermarking for images. *IEEE Trans. on Image Processing*, 2001, 10(5): 767–782.

[6] Pereira S, Ruanaidh J, Deguillaume F, Cscurka G, Pun T. Template based recovery of Fourier-based watermarks using log-polar and log-log maps. Proc. Int. Conference on Multimedia Computing and Systems, Florence, Italy. June 1999, (1): 870–874.

[7] Kutter M. Watermarking resisting to translation, rotation and scaling. Proc. SPIE Int. Symp. on Voice, Video, and Data Communication, November 1998, 3528: 423–431.

[8] Voloshynovskiy S, Deguillaume F, Pun T. Content adaptive watermarking based on a stochastic multiresolution image modeling. EUSIPCO2000, European Signal Processing Conference, Tampere, Finland, September 2000.

[9] Voloshynovskiy S, Herrigel A, Baumgaertner N, Pun T. A stochastic approach to content adaptive digital image watermarking. in Third International Workshop on Information Hiding, Dresden, Germany, September 29-October 1st, 1999, LNCS 1768; 270–285.

[10] E fstratiadis S, Katsaggelos A. Adaptive iterative image restoration with reduced computational load. *Optical Engineering*, 1990, 29(12): 1458–1468.

桑茂栋：　男，1980 年生，硕士生，研究方向为数字图像水印.

赵　耀：　男，1967 年生，教授，博士生导师，研究领域为图像编码、数字水印、基于内容的图像与视频检索、复杂网络环境下的可伸缩鲁棒视频编码研究、多媒体信息处理等.