数字水印的去同步攻击及其对策

李昌利 卢朝阳
(西安电子科技大学综合业务网国家重点实验室,西安 710071)

摘 要 有望解决数字媒体版权纠纷的数字水印技术日渐成熟,新算法不断涌现,面临的攻击也与日俱增。去同步攻击通过几何变形使得相关检测器失效,是一个很难应付的攻击。该攻击主要包括全局仿射变换攻击和局域任意扭曲攻击两种。有很多算法能很好对付前者,但后者几乎使现存的所有算法失效。本文对各种对付去同步攻击的对策分类加以综述,并分析了各自的不足,最后指出如何有效地抵御局域任意扭曲攻击依然是一个悬而未决的课题,尚值得深入研究。

关键词 数字水印 去同步攻击 仿射变换攻击 局域任意扭曲攻击 RST 不变性
中图法分类号: TP309.7 **文献标识码**: A **文章编号**: 1006-8961(2005)04-0403-07

Desynchronization Attacks on Digital Watermarks and Their Countermeasures

LI Chang-li, LU Zhao-yang
(*National Key Laboratory of Integrated Service Networks, Xidian University, Xi'an* 710071)

Abstract As a promising means to ascertain the copyright infringement of digital media, digital watermarking technique is getting more and more mature and new algorithms have been emerging. In the meantime, more and more attacks are confronting with it. This paper discusses desynchronization attack in detail, which disables the correlator through geometrical distortions and is very difficult to tackle. It mainly comprises global affine transformation attack and local random bending attack. Many algorithms can well withstand the former, but the latter almost defeats all existing algorithms. We categorize and summarize the countermeasures in detail, and analyze their deficiencies respectively. Finally we point out that how to effectively cope with desynchronization attack is still an open question and deserves a thorough research.

Keywords digital watermarking, desynchronization attack, affine transformation attack, local random bending attack, RST invariant

1 数字水印的去同步攻击

通信网络和图像处理技术的发展使得媒体以数字形式更方便地表征、存储、获取和分发,同时使得人们几乎不用付出任何代价就可非法拷贝和处理数字媒体。如何有效保护数字媒体所有者的合法权益成为了一个迫切需要解决的问题,数字水印技术应运而生,给业界带来了希望。在要保护的数字媒体中嵌入代表版权信息的水印信号,发现侵权行为时,提取相应的水印信号作为可靠的证据,有效解决了版权纠纷。

数字图像水印算法可谓层出不穷[1]。作为版权保护"最后一道防线"的图像不可见水印是最主要的研究课题。水印需要满足两个基本要求:不可见性(imperceptibility)和健壮性(robustness)。不可见性要求原始图像在嵌入水印后,仍然具有很好的视觉效果和商业价值。健壮性不仅要求嵌入水印的图像在经过合理处理(图像压缩、滤波、噪声污染、D/A 和 A/D 变换、再抽样、再量化和图像增强等)和常见的几何变换(旋转、平移、缩放和纵横比改变等)后仍能可靠提取出水印,而且要能有效对付各

基金项目:国家自然科学基金项目(60472083)
收稿日期:2003-11-11;**改回日期**:2004-09-07
第一作者简介:李昌利(1976~),男。2004 年于西安电子科技大学通信工程学院获通信与信息系统专业硕士学位,现于湛江海洋大学信息学院任教。研究兴趣为数字水印及信息论在其中的应用。E-mail:chlee_li@sohu.com

种恶意攻击。通过考虑人类视觉系统(HVS,human visual system),采用适当的视觉模型[2,3],在保证不可见性的前提下,尽可能提高水印信号能量,可以获得好的健壮性。

几何变换攻击虽然简单,但很具"杀伤力"。对整幅图像像素的空间移动同样改变了水印的物理位置,使得它的采样值不再在期望的位置,这与数字通信系统失去同步很类似。数字水印系统的"去同步"攻击由此而来。最基本的去同步攻击有全局旋转、平移、纵横比改变,它们可统一归结为仿射变换[4]:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{1}$$

其中,(x,y)和(x',y')分别为变换前后同一像素的坐标。显然只要确定了(a,b,c,d,e,f)这6个系数,再对遭受仿射变换攻击的图像进行相应的反变换即可建立同步。StirMark 软件包是一个很好的攻击工具,它包含一种称为局域任意扭曲攻击(RBA,random bending attack)的攻击[5]。RBA 造成的图像局域失真很小而最小均方误差很大,这使得人的视觉系统无法感知;然而它引入了微小的变形,使得提取出的水印信号与原始的信号错位了而不同步,并且不能用有限个独立参数描述,所以不像全局仿射变换那样易于处理。此外行列去除、剪切(crop)及抖动(jitter)等等亦属于去同步攻击的范畴。

现存的算法基本上都能抵御通常的图像处理操作,特别设计的算法能对付常见的几何操作,但是 RBA 攻击依然是数字水印系统面临的严峻挑战,它几乎使所有算法失效[6-8]。

以下针对静止图像的去同步攻击策略分类加以综述,并讨论了各自的优缺点。

2 利用不变性

2.1 利用 RST 不变性

在具有 RST(rotation,scaling and translation)不变性的变换域嵌入水印,这样即便图像遭受了 RST 变换,在变换域依然能检测到水印。

离散傅里叶变换(DFT)有以下特性:空间域的旋转导致频率域同样角度的旋转;空间域的循环移位不影响 DFT 的幅度;空间域的缩小(放大)导致频率域同样的放大(缩小)。

对笛卡儿坐标进行下式定义的 LPM(Log-polar mapping)映射变换:

$$x = e^{\mu}\cos(\theta),\ y = e^{\mu}\sin(\theta) \tag{2}$$

其中,(x,y)为笛卡儿坐标系的坐标表示,(μ,θ)为 LPM 域的坐标表示。

LPM 变换有以下特征:

$$(\rho x,\rho y) \leftrightarrow (\mu + \ln\rho,\theta) \tag{3}$$

$$(x\cos(\delta) - y\sin(\delta), x\sin(\delta) + y\cos(\delta)) \leftrightarrow (\mu,\theta + \delta) \tag{4}$$

其中,ρ 为缩放因子,δ 为旋转角度。

式(3)表明笛卡儿坐标系内的缩放(大小为 ρ)只导致 LPM 域 ρ 轴的平移(大小为 $\ln\rho$);式(4)表明笛卡儿坐标系的旋转(旋转角度 δ)只导致 LPM 域 θ 轴同样角度的平移。

FMT 变换(Fourier-Mellin Transform)过程为:先进行 LPM 变换,接着进行快速傅里叶变换(FFT,fast fourier transform)。笛卡儿坐标系内的缩放和旋转只导致 LPM 域的平移,而 FFT 变换具有平移不变性,所以 FMT 变换具有缩放和旋转不变性。

文献[9]提出了以下的水印嵌入算法:对原始图像进行 FFT 变换,接着进行 FMT 变换(即 LPM 变换之后再进行 FFT 变换),很显然,这样得到的变换系数具有 RST 不变性,从中选择一些系数嵌入水印信号;之后进行反 FFT 变换,反 FMT 变换(反 LPM 映射之后进行反 FFT 变换)。但是由于 LPM 变换在 FFT 域进行,这就需要在很广泛的动态范围内进行邻域系数的插值运算,算法实现起来很困难,而且图像有很明显的振铃效应;同时算法需要进行两次 FFT 运算,运算量很大。此外水印嵌入在两次 FFT 的幅度谱内,使得嵌入的频率可能不在原始图像的中频带,这样很难达到水印健壮性和不可见性的折衷。

文献[10]提出了一种 DFT 域的循环对称水印嵌入策略。频率在 R_1 与 R_2 间的同心圆被均分成 S 个扇区,同一个圆环上的同一个扇区嵌入相同的水印。这样构造的水印使得图像在$[-(\pi/s),+(\pi/s)]$范围内旋转任意角度都不影响水印的提取;类似地,缩放因子不影响归一化相关检测器的输出;平移不变性由 DFT 保证。

文献[11]提出了一种类似的、宜于实现的算法。记

$$i'(x,y) = i(\sigma(x\cos(\alpha) + y\sin(\alpha)) - x_0,\ \sigma(-x\sin(\alpha) + y\cos(\alpha)) - y_0) \tag{5}$$

为原始图像 $i(x,y)$ 经旋转、缩放及平移后的图像。式中 α、σ 及(x_0,y_0)分别为旋转、缩放及平移参数。

经过下式所定义的坐标转换

$$g(\theta) = \sum_j \log(I(\rho_j,\theta)) \quad \theta \in [0°,180°] \quad (6)$$

并记

$$g_1(\theta) = g(\theta) + g(\theta+90°) \quad \theta \in [0°,90°] \quad (7)$$

$i'(x,y)$的傅里叶幅度谱成为

$$|I'(\rho,\theta)| = |\sigma|^{-2} I(\rho - \log\sigma, \theta - \alpha) \quad (8)$$

显然 $g_1(\theta)$具有平移及缩放不变性,旋转导致 $g_1(\theta)$的循环移位,但把 θ 量化至最近的角度,则只有 90°移位,通过穷尽搜索可以处理它。不足之处是:式(7)中的求和运算,使得漏检概率很大;穷尽搜索很耗时;没有很好地处理小角度的旋转;没能解决好频率局域化问题。

为了克服以上不足之处,文献[12]通过求取图像不变质心(IC, invariant centroid)来获得平移不变性,在此基础上进行 LPM 变换即可获得 RST 不变性。通过迭代的方法可以得到一个几何攻击(包括剪切)或波形攻击对其没什么影响的点。由于 LPM 是在空间域进行的,避免了插值;由于只需要进行一次 FFT 运算,运算效率较高;由于只需要对水印信号进行反 LMP 变换再叠加在原始图像上得到最终图像,图像质量得以保证。

文献[13]提出了另外一种改进的 LPM 域算法。由傅里叶变换的平移不变性知 I'及 I 的傅里叶变换有以下关系

$$F'(w_\rho, w_\theta) = |\sigma|^{-2} e^{-j(w_\rho \cdot \log\sigma + w_\theta \cdot \alpha)} F(w_\rho, w_\theta) \quad (9)$$

幅度谱有以下关系

$$|F'(w_\rho, w_\theta)| = |\sigma|^{-2} |F(w_\rho, w_\theta)| \quad (10)$$

计算 F'和 F 的互功率谱

$$C = \frac{F'(w_\rho,w_\theta)F^*(w_\rho,w_\theta)}{|F'(w_\rho,w_\theta)F^*(w_\rho,w_\theta)|} = e^{j(w_\rho \cdot \sigma + w_\theta \cdot \alpha)} \quad (11)$$

再计算互功率谱相位的反傅里叶变换

$$D = IDFT(angle(C)) \quad (12)$$

其中,$IDFT$ 为反傅里叶变换,$angle(\,)$为求相位(角度)。

由于 $\delta(x-d)$的傅里叶变换为 e^{-jwd},所以式(12)给出了一个 2 维 δ 函数。在 LPM 域嵌入水印,LPM 具有平移不变性,由式(5)知空间域的旋转和缩放导致 LPM 域 θ 及 ρ 的移位。对嵌入水印的图像(遭受旋转、平移或缩放攻击后)和原始图像在 LPM 域依式(12)找出峰值点,通过这两个点确定平移参数来修正 LPM 幅度谱的 θ 及 ρ,再计算相关值进行阈值判断。用近似的反 LPM 代替反 LPM,消除了它带来的插值误差。此外文中采用了一个有效的、经验的视觉模型把水印嵌入在中频带。仿真结果表明该算法对旋转、平移及合理的缩放有很好的不变性。不足之处在于需要原始的、未嵌入水印的图像通过式(12)来确定 LPM 域 θ 及 ρ 的移位。

记实信号 $x(n)$的 DFT 变换为 $X(f)$,则其双谱(bi-spectrum)定义如下:

$$B(f_1,f_2) = X(f_1)X(f_2)X^*(f_1+f_2) \quad (13)$$

式中,$0 \leqslant f_2 \leqslant f_1 \leqslant f_1 + f_2 \leqslant 1$。

2 维信号 $f(x,y)$的 Radon 变换定义如下:

$$g(s,\theta) = \int_{-\infty}^{+\infty} f(x,\theta + sx)\,dx \quad (14)$$

定义

$$p(\theta) = \angle\left[\int_{f=0}^{0.5} B(f,f)\,df\right] = \angle\left[\int_{f=0}^{0.5} I^2(f,\theta)I^*(2f,\theta)\,df\right] \quad (15)$$

上式中的运算符"∠"指取相位。

显然积分双谱的相位具有平移和缩放不变性;Radon 变换具有旋转不变性。构造矢量 $\boldsymbol{P} = (p(\theta_1), p(\theta_2), \cdots, p(\theta_N))$,则 $\boldsymbol{P}$ 具有 RST 不变性。文献[14]提出通过改变 $\boldsymbol{P}$ 的值以嵌入水印的策略。较基于 FMT 的算法,该算法有以下优点:因为在计算 $p(\theta)$时避免了插值运算,在水印嵌入过程中混迭较少;因为水印嵌入在傅里叶相位谱,篡改(tamper)攻击较难。

文献[15]提出了一种简单获得 RST 不变性的方法。文中引入一种所谓的 LRHF(logarithmic radial harmonic function)振荡的同类模式,该模式具有很好的相关性、正交性和扩频特性。在原始图像中依像素叠加这些模式,即可得到嵌入水印的图像;通过在待检测图像中直接检测这些模式即可完成水印的检测。该算法利用信号的相位信息从而克服了插值问题。

2.2　利用矩不变性

$(\bar{x},\bar{y})$为图像 $f(x,y)$的质心,则中心矩定义如下:

$$\mu_{p,q} = \iint_I (x-\bar{x})^p (y-\bar{y})^q f(x,y)\,dxdy \quad (16)$$

归一化的中心矩定义如下:

$$\eta_{p,q} = \frac{\mu_{p,q}}{(\mu_{0,0})^\gamma} \quad (17)$$

其中,$\gamma = \dfrac{(p+q+2)}{2}$。

有两组矩不变量集分别具有仿射不变性和正交变换不变性。这里只给出具有仿射不变性的一组矩 $\phi_i(i=1,\cdots,7)$ 中的一个 ϕ_1

$$\phi_1 = \eta_{0,2} + \eta_{2,0} \tag{18}$$

令

$$\phi_i^* = |\log_{10}\phi_i| \tag{19}$$

文献[16]提出以下算法:预先定义函数 f, 对原始图像 I 施加扰动 ΔI 得到嵌入水印的图像 $\bar{I}$ 满足:

$$v-\varepsilon \leqslant f(\boldsymbol{\Phi}^*) \leqslant v+\varepsilon \tag{20}$$

式中,$\boldsymbol{\Phi}^*$ 是包含图像 $\bar{I}$ 几何矩 $\phi_i(i=1,\cdots,7)$ 的矢量;ε 是由实验确定的容忍度,取值很小。在水印检测阶段,若待检测图像满足式(20),则认为水印存在。该算法收敛性较好,不足之处是图像保真度较差,嵌入水印前后的图像对比度变化较大;不能容忍任何的纵横比变化及剪切操作。

文献[17]提出了改进的算法,使得在水印嵌入过程中图像没有很明显的对比度变化。仿真结果表明此算法能抗缩放、纵横比变化、行列去除及通常的反转、平移及旋转(没有剪切)等几何变换攻击。

2.3 利用 DT 三角形不变性

基于内容的水印算法提供了一种新颖的解决方法。文献[18]把水印与图像表征有机结合起来,利用特征点检测算子 Harris 算子检测出图像的局部极大点,并对这些极大点进行德劳内分割(DT, Delaunay Tessellation)以形成图像的三角形表征。嵌入水印的图像在遭受几何变换攻击后,有一些三角形会保持原来的物理位置和形状,水印检测在其中进行。水印嵌入过程如下:(1)用 Harris 算子对原始图像进行处理,得到局部极大点集;(2)对点集进行 DT,水印嵌入到这些不相交的三角形内。对待检测图像进行同样的 Harris 检测和 DT,计算那些保持原来顶点坐标位置和形状的三角形与相应的水印之相关值,把它们累加起来进行门限判决。该算法有很好的健壮性,几乎能抗目前已知的所有攻击。然而算法的有效性有赖于,图像特别是高纹理区在遭受几何攻击后,Harris 算子提取出的特征点 DT 后的三角形能否保持原来的物理位置和形状。

2.4 基于图像特征不变性

文献[19]中分析了基于图像特征不变性的水印算法。算法框架包括以下 3 个步骤:

(1) 特征点提取　提取出在几何攻击下有很好健壮性的特征点,有基于分割的提取方法及基于边缘的提取方法。前者又有基于 Gibbs 随机场和局域算子的分割法、基于 K-均值颜色聚类的分割法。

(2) 登记成标准几何形式　用上一步提取出的特征点集把图像分割成基本的小块,再把这些小块变换成标准几何形式,最后在这种标准形式中嵌入水印。接着对其进行反变换得到嵌入水印后的图像。尽管标准几何形式可以是任何形状,但选择三角形或四边形更好,因为它们分别与 6 点仿射变换及 8 点透射变换相对应。

(3)水印的嵌入或提取　水印的嵌入过程如上所述,水印的提取过程与常见的算法一样。

显然这类算法的健壮性依赖于特征点提取算子的可重复性及准确性。

3 求出几何变换系数,再进行反变换

若能获得图像遭受的几何变换系数,再进行相应的反变换,即可获得同步。利用规则点集的一些先验知识可以对仿射变换系数进行估计,它们可以是函数的局部极大点或峰值点。信号的自相关函数或周期信号的傅里叶变换幅度谱可以产生这样的点集,它们规则地排列或等间隔分布。下面的 3.1 节和 3.2 节中的算法都是利用图像遭受仿射变换前后的峰值点通过式(1)求出仿射变换系数。不同之处在于 3.1 节算法中的峰值点由特殊结构的水印信号自身产生,而 3.2 节由嵌入的参考水印信号产生。3.3 节是另外一种估计缩放因子及旋转角度的算法。

3.1 运用周期系列作为水印信号

文献[20]提出了这样一种算法:(1)把具有周期模式的水印重复多次嵌入图像中;(2)对待检测的图像进行频谱分析,对峰值点进行 Hough 变换,估计仿射变换矩阵;(3)对待检测图像进行相应的反变换,提取水印。该算法引入了健壮的最大似然惩罚函数从所有可能的矩阵中搜索最优解。

3.2 同时嵌入用于同步的模板

在原始图像中同时嵌入提供信息的水印信号(informative watermark)和所谓的参考水印(reference watermark)。前者包含真正的水印信息,后者只是用于同步的模板。

文献[21]的算法在原始图像中嵌入一个自相关函数有 9 个峰值点的参考水印,用以估计几何攻击的参数。算法固有的缺点是:非中心的 8 个峰值点对攻击很脆弱。

文献[22]提出一个更有效的算法。通过周期性地嵌入块状模板,使得傅里叶幅度谱的峰值点分布在整幅图像。

文献[23]中先对原始图像进行 DFT 变换,水印信号嵌入在 f_1 和 f_2 之间的中频带上以加强健壮性。在频率域,两条过原点倾角在 θ_1 在 θ_2 之间,距离原点在 R_1 和 R_2 之间的直线上各嵌入 7 个点形成模板。在经受线性变换后,这两条直线被变换成另外两条过原点的直线。仿真结果表明算法对仿射变换有较好的免疫力。

文献[24]提出了另外一种算法。利用 DFT 的平移不变性,在 DFT 域嵌入提供信息的水印信号;在空间域直接重复嵌入一个小方形模板。在检测阶段,把待测试图像经过维纳滤波器处理,所得图像的自相关函数会出现峰值点。通过与嵌入了水印而没有遭受旋转和缩放的图像自相关函数峰值点进行比较即可确定变换矩阵,对待测试图像进行同样的反变换,即可建立同步。

3.3　通过其他方法求出几何变换系数

文献[25]提出了一种通过求 ESDR(edges standard deviation ratio)来估计图像遭受的缩放;通过求 AEAD(average edges angels difference)来估计图像遭受的旋转。对图像 $f(x)$ 进行非正交的离散小波变换,取模 $|Wf(s,x)|$,满足下式的点为极大值点:

$$|Wf(s,x)| \leqslant |Wf(s,x^*)| \tag{21}$$

把缩放因子为 s 的图像中满足式(21)的所有 N 个点连接在一起,得到极大值点的轮廓。下式给出 ESD(edges standard deviation):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_0)^2 + (y_i - y_0)^2} \tag{22}$$

其中,$x_0 = \dfrac{\sum_i x_i}{N}$ 和 $y_0 = \dfrac{\sum_i y_i}{N}$ 为坐标均值。

若 σ_s 及 σ_0 为图像缩放前后的 ESD,ESDR 定义为

$$\gamma = \sigma_s / \sigma_0 \tag{23}$$

具体实现时,先对所有的极大值进行归一化处理,记归一化极大值为 $WM(x_i, y_i)$,若它大于阈值 T,则其为极大值点。依式(23)对由不同阈值 T 得到的轮廓求得相应的 ESDR,再取均值作为缩放因子的估计值。仿真结果表明得到的估计值精确度很高,而且对非对称剪切有很好的健壮性。类似可以定义 AEAD。

4　运动估计补偿的方法

几何失真可以看作失真图像(嵌入水印的图像遭受几何攻击)和参考图像(原始图像或没有遭受攻击的嵌入了水印的图像)之间的某种运动。通过运动估计补偿,可以把遭受局域 RBA 攻击的图像进行一定的几何校正,有助于水印的检测。

文献[8]提出了一种可变形网格基的方法。失真图像和参考图像用相同的网格表征;通过使两者间的匹配误差极小化来估计出失真场;最后用失真场来校正失真图像。在校正后的图像中即可用通常的方法检测出水印。

文献[26]中通过求出失真图像和参考图像之间的运动矢量进行几何校正。算法包括运动估计和运动补偿两个过程。首先对失真图像和参考图像进行 $m_{\max}$ 级复小波变换(CWT, complex wavelet tansform)。校正由 m 级到第 1 级的顺序分级进行。记 m 级 s 子带位置 n 处的 CWT 系数为 $D^{(m,s)}(n)$,失真图像较 n 位移为 $f=(f_x,f_y)$ 处对应系数为 $D^{(m,s)}(n+f)$,插值过程为

$$D^{(m,s)}(n+f) \approx \sum W_f^{(m,s)}(k) D^{(m,s)}(n+f) \tag{24}$$

式中,W 为核函数。定义子带平方差 $SD^{(m,s)}$(subband squared difference)为

$$SD^{(m,s)}(n,f) = |D_1^{(m,s)}(n+f) - D_2^{(m,s)}(n)|^2 \tag{25}$$

类似地定义累积平方差别 $CSD^{(m)}$(cumulative squared difference)为

$$CSD^{(m)}(n,f) = \begin{cases} CSD'^{(m+1)}(n,f) + SD^{(m)}(n,f) & m < m_{\max} \\ SD^{(m)}(n,f) & m = m_{\max} \end{cases} \tag{26}$$

其中,$CSD'^{(m+1)}(n,f)$ 是 $CSD^{(m)}(n,f)$ 的双线性插值。

$CSD^{(m)}$ 的极小给出了 m 级的运动估计。通过对 m 级错误运动矢量的处理,得到 $m-1$ 级更精细的初始运动估计。最终每个 4×4 像素块得到一个运动矢量,在这样大小的尺度上,几何失真可以近似为运动矢量的简单平移,只需把每个像素块移回原来的位置即可。仿真结果表明该算法在图像遭受小的几何失真攻击时,有助于水印的检测;然而一旦面临大的几何失真攻击,如旋转、缩放或剪切时,则无能为力。

5　图像归一化的方法

图像归一化是指通过几何变换把图像变成一种

标准形式。对原始图像和待检测图像进行同样的归一化处理,水印的嵌入和检测在归一化的图像中进行。这种方法提供了另外一种简单的能抗几何攻击的策略。

文献[4]中对质心为(g_x,g_y)的图像$f(x,y)$定义变量 *OII*(orientation indicator index):

$$OII_f=\sqrt{\mu_x^2+\mu_y^2} \tag{27}$$

其中,$\mu_x=\int_R(x-g_x)f(x,y)\,dxdy$

$$\mu_y=\int_R(y-g_y)f(x,y)\,dxdy$$

定义

$$\alpha=|\max_x\{x;f(x,y)\neq 0\}-\min_x\{x;f(x,y)\neq 0\}| \tag{28}$$

$$\beta=|\max_y\{y;f(x,y)\neq 0\}-\min_y\{y;f(x,y)\neq 0\}| \tag{29}$$

$f(x_d,y_d)$为$f(x,y)$经过旋转θ和缩放$\{1/\alpha,1/\beta\}$后的图像,其坐标为

$$\begin{bmatrix}x_d\\y_d\end{bmatrix}=\begin{bmatrix}1/\alpha & 0\\0 & 1/\beta\end{bmatrix}\begin{bmatrix}\cos(\theta) & \sin(\theta)\\-\sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} \tag{30}$$

图像归一化过程为

(1)把原始图像的原点重新定位在(g_x,g_y);计算$\{\alpha;\beta\}$,对一定的θ,依式(30)计算(x_d,y_d);

(2)对$[0,\pi/2]$区间的θ,反复迭代进行上一步,有最大 *OII* 值的图像即为最终的归一化图像$f^{\theta}(x,y)$。仿真结果表明该算法有较好的抗仿射变换性。

文献[27]提出了另外一种基于几何矩的归一化方法。对原始图像和待检测图像,分别计算一些与图像本身的几何矩有关的参数来对图像进行归一化处理。仿真结果表明算法能抗旋转、缩放及翻转(flipping)攻击,通过归一化图像的兴趣区域(ROI, region of interest)还可以抗剪切攻击。

6 信道模型估计法

众所周知,数字水印可以看作是一个隐蔽信道的通信过程,各种攻击可视为信道噪声,通过模型化信道噪声,可以设计健壮的算法。

文献[28]通过把水印信道模型化为加性噪声和几何噪声组成的复合信道,研究了抖动(jitter)攻击,引入索引变量的概念。在抖动信道中,索引变量形成一个马尔可夫链,这使得设计一个有效的、适合信道的算法成为可能。通过调整马尔可夫链(场)参数,即状态间的转移概率,可以较好地模型化各种局域几何攻击。

文献[29]提出了一种针对所谓的 SCS(scalar costa scheme)水印算法的信道用以模拟水印系统局域的失同步。文中指出不精确的同步导致水印系统的码间干扰,并给出了在此条件下系统的最大水印容量,提出一种利用参考水印估计采样格子遭受的几何变形的策略。

文献[7]把信道模拟为平稳攻击信道$A(y|z)$和确定性的失同步信道$W(x,\theta)$的级联$A\circ W$,θ代表旋转、平移及缩放等几何参数。通过最优化方法得到θ,即可设计用于同步的模式。

7 结 语

在具有 RST 不变性的变换域,对连续信号来说,即便遭受了仿射变换攻击,不变性也能保持;然而对离散的图像信号来说,即使它没有遭受任何几何变换,由于空间域和变换域之间没有精确的双向映射,需要进行某种形式的内插运算,从而带来了插值误差,降低了图像质量。尽管这些算法能对付仿射变换攻击,但由于 FFT 对剪切很敏感,所以它们抗剪切性能很差[25]。此外由于在这些变换域缺乏相应的心理视觉模型,很难使水印的嵌入强度在不可见性和健壮性间达到一个很好的折衷;算法的有效性也缺乏理论解释[12]。

利用周期模式或参考水印的自相关函数或幅度谱的峰值点来确定变换矩阵的方法,由于这些峰值点宜于发现,所以也宜于去除。文献[30]提出一种通过邻域平滑的方法去除这些峰值点的模板攻击,文献[31]详尽地分析了这类算法,并进行了大量的仿真实验,证实这类算法易于攻击。此外嵌入模板作为参考水印的方法需要同时嵌入两个水印信号,降低了图像质量和水印容量;由于所有采用该算法的图像共用一个模板水印/记录水印(registration watermark),共谋攻击很容易得逞[11]。

只有文献[18]、[26]、[28]等几种算法能对付局域几何变换攻击。然而文献[18]中的算法需要提取出的特征点有较好的稳定性;文献[26]中的算法不能对付全局仿射变换攻击;文献[28]提出用马尔可夫链(场)来模型化几何攻击的方法,是一个很好的尝试。

越来越多的攻击方法使得人们对数字水印技术充满忧虑甚至失望。去同步攻击就是其中一个很致

命的攻击,它造成的图像失真很小,然而几乎使现存的所有算法失效。本文对文献中对付去同步攻击的策略加以分类介绍并分析了各自的优缺点。有较多的算法能很好对付全局仿射变换攻击,然而大多数算法在局域任意扭曲攻击面前显得无能为力,因此如何有效地对付它依然是水印技术真正走向实用的不可回避的难题,值得深入研究。

参考文献(References)

1 Hartung F, Kutter M. Multimedia watermarking techniques [J]. Proceedings of the IEEE, 1999, **87**(7):1079 ~ 1107.

2 Podilchuk C L, Zeng W. Image-adaptive watermarking using visual models [J]. IEEE Journal on Selected Areas of Communications, 1998, **16**(4):525 ~ 539.

3 Kutter M, Winkler S. A vision-based masking model for spread-spectrum image watermarking [J]. IEEE Transactions on Image Processing, 2002, **11**(1):16 ~ 25.

4 Dong P, Galatsanos N P. Affine transformation resistant watermarking based on images normalization [A]. In: Proceedings of International Conference on Image Processing 2002 [C], Rochester, USA: 2002: 489 ~ 492.

5 Petitcolas F A P. StirMark [EB/OL]. http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/, 1999-03-01.

6 Voloshynovskiy S, Deguillaume F, Pun T. Multibit watermarking robust against local nonlinear geometrical distortions [A]. In: Proceedings of International Conference on Image Processing 2001 [C], Thessaloniki, Greece, 2001:999 ~ 1002.

7 Moulin P, Ivanović A. The fisher information game for optimal design of synchronization patterns in blind watermarking [A]. In: Proceedings of International Conference on Image Processing 2001 [C], Thessaloniki, Greece, 2001:550 ~ 553.

8 Dong P, Brankov J. Geometric robust watermarking through mesh based correction [A]. In: Proceedings of International Conference on Image Processing 2002 [C], Rochester, USA, 2002:493 ~ 496.

9 Joseph J K. Ruanaidh Ó, Pun T. Rotation, scale and translation invariant spread spectrum digital image watermarking [J]. Signal Processing, 1998, **66**(3):303 ~ 317.

10 Solachidis V, Pitas I. Circularly symmetric watermark embedding in 2-D DFT domain [J]. IEEE Transactions on Image Processing, 2001, **10**(11):1741 ~ 1753.

11 Lin C Y, Wu M, Bloom J A, *et al.* Rotation, scale and translation resilient watermarking for images [J]. IEEE Transactions on Image Processing, 2001, **10**(5):767 ~ 782.

12 Kim B S, Choi J G, Park C H, *et al.* Robust digital image watermarking method against geometrical attacks [J]. Real-time Imaging, 2003, **9**(2): 139 ~ 149.

13 Zheng D, Zhao J, Saddik A E. RST invariant digital image watermarking based on log-polar mapping and phase correlation [J]. IEEE Transactions on Circuits and System for Video Technology, 2003, **13**(8): 753 ~ 765.

14 Kim H S, Baek Y, Lee H K. Rotation, scale, and translation invariant watermark using higher order spectra [J]. Optical Engineering, 2003, **42**(2):340 ~ 349.

15 Fletcher P A, Larkin K G. Direct embedding and detection of RST invariant watermarks [A]. In: Proceedings of Information Hiding 2002 [C], Noordwijkerhout, The Netherlands, 2002:129 ~ 144.

16 Alghoniemy M, Tewfik H A. Image watermarking by moment invariant [A]. In: Proceedings IEEE International Conference on Image Processing 2000 [C], Vancouver, 2000:73 ~ 76.

17 Lu C S. Towards robust image watermarking: combining content-dependent key, monment normalization, and side-informed embedding [J]. Signal Processing: Image Communication, 2005, **2**(20):129 ~ 150.

18 Bas P, Chassery J-M, Macq B. Geometrically invariant watermarking using feature points [J]. IEEE Transactions on Image Processing, 2002, **11**(8):825 ~ 837.

19 Celik M U, Saber E, Sharma G, *et al.* Analysis of feature-based geometry invariant watermarking [A]. In: Proceedings of SPIE: Security and Watermarking of Multimedia Contents III [C], San Jose, USA, 2001, **4314**: 261 ~ 268.

20 Deguillaume F, Voloshynovskiy S, Pun T. A method for the estimation and recovering from general affine transforms in digital watermarking applications [A]. In: Proceedings of SPIE [C], San Jose, USA, 2002, **4675**:313 ~ 322.

21 Kutter M. Watermarking resistant to translation, rotation and scaling [A]. In: Proceedings of SPIE [C], Boston, USA: November, 1998, **3530**: 423 ~ 431.

22 Voloshynovskiy S, Deguillaume F, Pun T. Multi-bit digital watermarking robust against local nonlinear geometrical distortions [A]. In: Proceedings of IEEE International Conference on Image Processing 2001 [C], Greece, 2001, **10**:999 ~ 1002.

23 Pereia S, Pun T. Robust template matching for affine resistant image watermarks [J]. IEEE Transactions on Image Processing, 2000, **9**(6):1123 ~ 1129.

24 Po-Chyi Su, Kuo C C J. An image watermarking scheme to resist generalized geometrical transformations [A]. In: Proceedings of SPIE [C], Boston, MA, USA, 2001, **4209**:354 ~ 365.

25 Alghoniemy M, Tewfik A H. Geometric distortions correction in image watermarking [A]. In: Proceedings of SPIE 2000 [C], San Jose, USA, 2000, **3971**:82 ~ 89.

26 Loo P, Kingsbury N. Motion estimation based registration of geometrically distorted images for watermark recovery [A]. In: Security and Watermarking of Multimedia Contents, Part of SPIE Electronic Imaging [C], San Jose, 2001, **4314**:606 ~ 617.

27 Alghoniemy M, Tewfik A H. Geometric distortions correction through image normalization [A]. In: Proceedings of ICME [C], New York, USA, 2000:1291 ~ 1294.

28 Baudry S, Nguyen P, Maître H. Optimal decoding for watermarks subject to geometrical attacks [J]. Signal processing: Image Communication, 2003, **18**(4):297 ~ 307.

29 Bäuml R, Eggers J J, Tzschoppe R, *et al.* A channel model for watermarks subject to desynchronization attacks [A]. In: Proceedings of SPIE [C], San Jose, USA, Jan, 2002, **4675**:19 ~ 25.

30 Herrigel A, Voloshynovskiy S, Rytsar Y B. The watermark template attack [A]. In: Proceedings of SPIE [C], San Jose, USA, 2001, **4314**:4314 ~ 4346.

31 Rodríguez M Á, González F P. Analysis of pilot-based synchronization algorithms for watermarking of still images [J]. Signal Processing: Image Communication, 2002, **17**(8):611 ~ 633.