中国科学: 信息科学 2016 年 第 46 卷 第 5 期: 627–642
SCIENTIA SINICA Informationis
《中国科学》杂志社
SCIENCE CHINA PRESS

论 文

基于多方向差分的重采样取证技术

彭安杰[①②†], 曾辉[①③†], 康显桂[①③*]

① 中山大学数据科学与计算机学院, 广州 510006
② 西南科技大学计算机学院, 绵阳 621010
③ 广东省信息安全技术重点实验室, 广州 510006
* 通信作者. E-mail: isskxg@mail.sysu.edu.cn
† 同等贡献

收稿日期: 2015–07–02; 接受日期: 2015–09–24
国家自然科学基金 (批准号: 61379155, U1536204, 61332012, 61502547) 和广东省自然科学基金重点项目 (批准号: s20130200 12788) 资助

摘要 重采样操作常用于数字图像篡改, 重采样的盲取证受到了研究者的关注. 已有的重采样取证算法主要关注取证检测器的有效性, 而忽略了检测器的安全性, 即恶意攻击者对检测器的攻击. 目前, 针对重采样取证的攻击已出现. 该攻击使得基于周期性 "指纹" 的重采样取证检测器失效. 本文提出了一种基于多方向差分的重采样取证技术. 它同时考虑了重采样取证的有效性与安全性. 首先根据方向性和对称性将多方向差分分组, 然后分别建立自回归模型 (auto-regressive model, AR) 并提取出自回归系数和直方图特征, 最后将所有分组特征组合成重采样检测特征. 在由多个图像数据库组成的混合图像库上的测试结果表明所提出的算法既能有效地检测重采样操作, 也能检测已有的恶意攻击. 此外, 本文提出的算法也大幅度地提升了下采样操作的检测准确率以及针对 JPEG 压缩的鲁棒性. 篡改图像上的检测结果也证实所提算法兼顾了有效性与安全性.

关键词 数字图像取证 重采样 多方向差分 自回归模型

1 引言

数字图像取证 (digital image forensics) 是多媒体安全领域的一个重要研究方向. 不同于主动认证技术, 如数字水印 [1]、Hash 校验 [2], 数字图像取证是一种被动认证技术, 即不依赖任何预签名或预嵌入信息的前提下, 对数字图像的原始性、真实性及来源进行鉴别. 数字图像取证包括操作取证和篡改取证. 常见的操作取证有重采样取证 [3~10]、中值滤波取证 [11,12]、对比度增强取证 [13]、锐化取证 [14]、JPEG 压缩取证 [15,16] 等. 图像操作取证能还原图像的处理历史, 辅助鉴别图像的真实性和原始性. 篡改操作取证用于辨别图像内容的真伪, 它在遏制虚假新闻、维护司法公正等方面有着重要的意义 [17,18].

重采样 (re-sampling) 是图像几何变换 (缩放、旋转等) 中必需的操作. 它的取证具有重要的意义: (1) 还原图像的处理历史, 适用于对图像原始性要求高的场合. (2) 辅助判断图像内容的真实性. 在图

引用格式: 彭安杰, 曾辉, 康显桂. 基于多方向差分的重采样取证技术. 中国科学: 信息科学, 2016, 46: 627–642, doi: 10.1360/N112015-00034

© 2016《中国科学》杂志社

像的篡改过程中篡改者经常使用缩放操作或旋转. 取证重采样为检测篡改操作提供了辅助线索. (3) 提高隐蔽通信的安全性. 数字隐写术常将信息隐藏于常见的网络图片中. 互联网上流传的图片很多都经过了下采样操作. 文献 [19] 分析了各种下采样操作对隐写术安全性的影响, 并指出将信息隐藏于由双三次插值生成的下采样图片中具有安全隐患.

目前已有的重采样取证算法大致可以分为两类: (1) 基于周期性痕迹的取证 [3~7]; (2) 利用特征训练支持向量机 (SVM) 进行二分类的取证 [9,10]. 等距采样使得重采样操作后的图像留下了周期性的痕迹. 文献 [3] 观察到重采样图像的相邻像素之间存在周期性的依赖关系, 并用概率图去估计相邻像素间的线性依赖关系. 文献 [4,5] 利用二阶差分方差的周期性去检测重采样操作. Kirchner [6] 指出了基于概率图的取证算法 [3] 和基于差分方差周期性算法 [4,5] 之间的关系, 并提出了一种快速的检测算法. 上述基于周期性痕迹的检测算法在上采样操作的取证中取得了较好的效果, 但是在下采样图像和 JPEG 压缩图像上的检测效果较差. Feng 等 [9] 和 Li 等 [10] 分别从离散 Fourier 变换 (DFT) 域中提取能量特征和矩特征, 然后利用 SVM 训练生成分类器. 基于 SVM 的重采样检测算法能较好地识别上采样和下采样操作.

已有的重采样取证算法皆假设没有恶意的攻击者存在, 这种假设在实际应用中是不合理的. 反取证 (anti-forensics) 算法是一种专门攻击取证算法的技术. 它通过擦除取证技术依赖的 “指纹” 痕迹使得取证技术失效. Kirchner 等提出了一种攻击重采样取证的反取证算法 [20], 它通过在重采样点位置上添加 Gauss 扰动, 擦除周期性痕迹, 从而使得基于周期性的检测算法 [3~7] 失效.

反取证技术的出现导致了对取证技术的安全性研究 [21]. 安全的操作取证算法不仅能有效地辨识某种操作, 而且能抵抗反取证技术的攻击. 为此, 本文提出了一种基于多方向差分的安全重采样取证方法, 它能正确地区分原始图像和重采样图像 (包括正常重采样图像和经过反取证生成的重采样图像). 下文将未遭受反取证攻击的正常重采样图像称为重采样图像, 将经过反取证攻击的重采样图像称为反取证图像.

我们发现重采样图像和反取证图像皆通过插值生成, 因此这两类图像包含插值操作的痕迹. 本文利用插值操作的痕迹去检测重采样和反取证操作, 主要工作有: (1) 利用自回归系数去揭示原始图像、重采样图像和反取证图像在相邻像素间的不同依赖关系; (2) 利用多方向差分的直方图去揭示不同种类图像在高频域的统计特征. 实验结果表明本文提出的方法的检测错误率远低于已有的一些检测技术 [9,10], 并且所提方法对抗 JPEG 压缩的鲁棒性较强, 相对于现有方法有着重要的改善.

2 相关工作

2.1 重采样操作

图像的几何变换由两个基本操作组成: (1) 重采样, 也叫作坐标的空间变换; (2) 插值. 设原始图像为 $I(x, y)$, 经过几何变换后的图像为 $R(u, v)$, 它们尺寸大小分别为$r \times c$ 和 $m \times n$. 重采样步骤通过如下逆映射计算新像素 (u, v) 在原始图像坐标系 (x, y) 上的位置 (x', y'):

$$(x', y') = T^{-1}(u, v), \tag{1}$$

这里 T^{-1} 代表仿射变换的逆映射, 常见的尺度变换和旋转变换的仿射映射如下:

$$\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

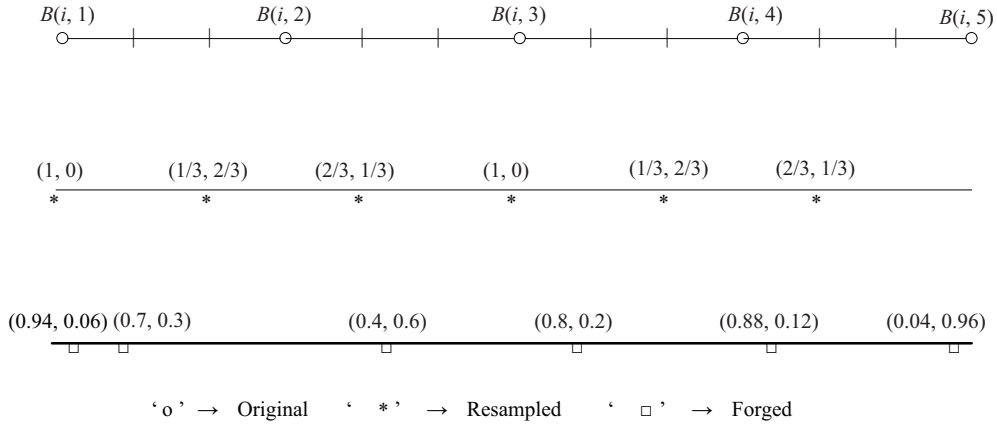

图 1 重采样插值过程. 第 1∼ 3 行分别为原始图像、重采样图像 ($s = \frac{3}{2}$, 双线性插值) 以及反取证图像在第 i 行的采样点及其插值权值

Figure 1 Example of the up-scaling ($s = \frac{3}{2}$, bilinear) process for the i-th row of image B. First line: original image; second line: up-scaled image; third line: up-scaled image undergone attack 1

其中 s 表示尺度变换因子, $s > 1$ 为上采样, $s < 1$ 为下采样, θ 表示旋转角度. (x', y') 的灰度值由邻近的原始图像 $I(x, y)$ 的像素插值形成.

图像尺度变换一般按照行和列分别进行变换. 变换后的图像可用下式表示:

$$R = A^{\mathrm{r}} I A^{\mathrm{c}} = B A^{\mathrm{c}}, \tag{2}$$

这里 $A^{\mathrm{r}}(m \times r)$ 和 $A^{\mathrm{c}}(c \times n)$ 分别表征了行和列的插值过程, $B(m \times c)$ 代表行插值之后形成的中间图像. 由于行列方向的重采样过程相同, 式 (2) 可以将二维重采样图像的分析简化为一维信号的重采样分析. 下文将只分析行方向上的重采样过程. 矩阵 A^{c} 由尺度因子 s 和插值核决定, 其中尺度因子决定重采样点在原始图像坐标系中的位置, 插值核决定邻近原始像素的选取以及相应的插值权值. 一般来说, 较近的原始像素被赋予更大的插值权值. 图 1 第 2 行展示了由中间图像 B (图 1 第 1 行) 的第 i 行插值生成重采样图像 R 的第 i 行的部分过程, 其他行的重采样过程也类似. 双线性插值函数的宽度为 2, 因此重采样元素由邻近的两个原始像素生成, 如 $R(i, 2) = \frac{1}{3}B(i, 1) + \frac{2}{3}B(i, 2)$. 图 1 中的重采样图像的矩阵 A^{c} 由式 (3) 表示:

$$A^{\mathrm{c}} = \begin{pmatrix} 1 & \frac{1}{3} & 0 & 0 & 0 \cdots \\ 0 & \frac{2}{3} & \frac{2}{3} & 0 & 0 \cdots \\ 0 & 0 & \frac{1}{3} & 1 & \frac{1}{3} \cdots \\ 0 & 0 & 0 & 0 & \frac{2}{3} \cdots \\ & & & \cdots & \end{pmatrix}. \tag{3}$$

2.2 重采样取证算法

Popescu 等利用重采样图像相邻像素间的周期相关性检测重采样操作 [3]. 等距采样使得插值权值序列呈现周期性, 如图 1 第 2 行所示. 这种插值权值序列使得 A^{c} 中的相邻列呈现出周期性的线性组合关系. 在式 (3) 所示的 A^{c} 中, 其列 $A^{\mathrm{c}}(:, k)$ 之间的周期性线性组合: $A^{\mathrm{c}}(:, 3k) = -\frac{1}{3}A^{\mathrm{c}}(:, 3k-2) + A^{\mathrm{c}}(:, 3k-1) + \frac{1}{3}A^{\mathrm{c}}(:, 3k+1)$ $(k = 1, 2, 3, \ldots)$. 由式 (2) 可知, 重采样后的像素值 $R(i, k) = B(i, :)A^{\mathrm{c}}(:, k)$,

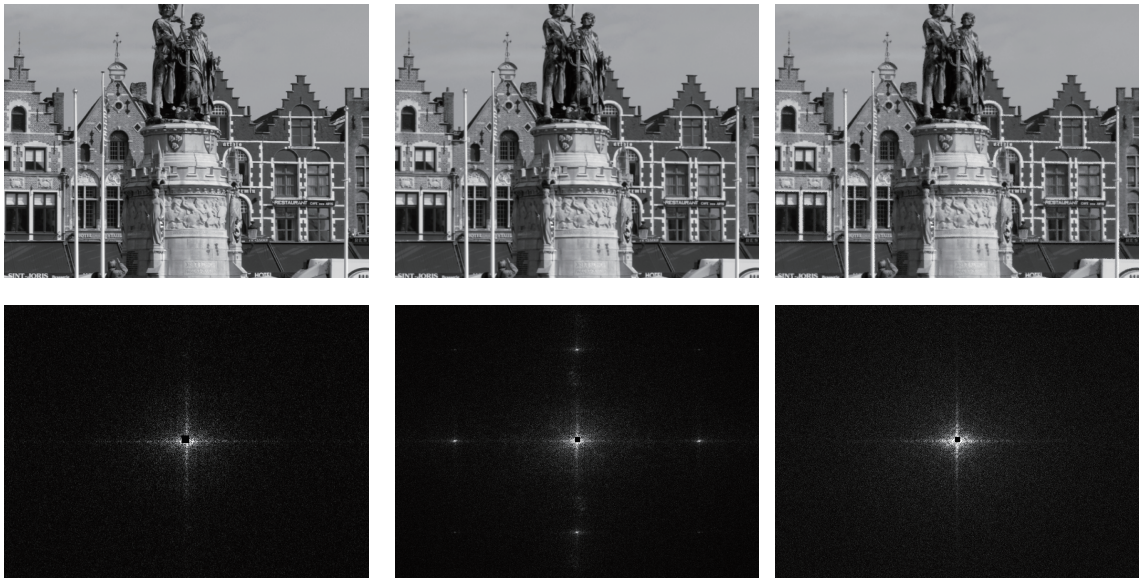

图 2　各类图像 (第 1 行) 及其概率图的 DFT 幅度谱 (第 2 行). 从左到右依次为原始图像、重采样图像 ($s=\frac{3}{2}$, 双线性插值)、反取证图像 (攻击 1, $s=\frac{3}{2}$, $\sigma=0.4$, 双线性插值)

Figure 2　Top row (from left to right): original image, up-scaled image, up-scaled image undergone attack 1; bottom row: the corresponding magnitude of the Fourier transform of p-map

因此 $R(i,3k)=-\frac{1}{3}R(i,3k-2)+R(i,3k-1)+\frac{1}{3}R(i,3k+1)$, 即重采样图像相邻像素间存在周期性的线性关系. 对于尺度因子 $s=\frac{a}{b}$ (a, b 互素), 上述线性关系的周期等于 a. 为了揭示重采样图像相邻像素间的周期性线性关系, Popescu 等利用最大期望值算法 (EM) 去估计一个概率图 (p-map) [3]. 该概率图反映了每个像素与其相邻像素线性相关的概率. 周期性的线性相关使得概率图也呈现出了周期性现象. 图 2 第 2 行显示上采样图像概率图的 Fourier 变换幅度谱中出现了周期性现象 (除中心点外的其余 4 个亮点). 该算法在未压缩的重采样图像的检测中取得了较好的效果, 但是该算法对抗 JPEG 压缩的鲁棒性不强. 原因在于: (1) JPEG 压缩削弱了 8×8 块内的重采样痕迹; (2) JPEG 压缩中使用的 8×8 分块也产生了周期性. 块效应产生的周期性干扰了重采样的周期性.

Li 等 [10] 从 DFT 域的幅度谱提取特征检测重采样. 具体做法是: 首先对图像进行 Laplace 高通滤波, 紧接着将滤波后图像的绝对值进行 DFT 变换, 然后将整个频率范围均分为 20 个子带 (Ω) 并在每个子带上提取矩特征 ($P(\Omega)$), 最后将所有子带的矩特征合并为 20 维的检测特征, 记为 F_{M}. 子带的矩特征由式 (4) 表示, 它反映了该子带的平均能量分布. 该算法在上采样图像的检测上取得了较好的效果, 但是对抗 JPEG 压缩的鲁棒性有待提高.

$$\begin{aligned}&\Omega=\{w<D_{u,v}\leqslant w+\Delta\},\\&P(\Omega)=\sum_{(u,v)\in\Omega}\frac{1}{500}\frac{|R(u,v)|}{E_d}D_{u,v},\end{aligned}\tag{4}$$

上式中, $D_{u,v}$ 表示 Chebyshev 距离, $|R(u,v)|$ 表示 DFT 幅度谱, E_d 表示所有 $|R(u,v)|$ (除直流分量外) 的平均值, Δ 表示子带之间的间隔.

2.3 针对重采样取证的反取证算法

Kirchner 等提出了一种反取证技术 [20], 它主要攻击基于周期性的重采样取证技术 [3~7]. 由 2.2 小节的分析可知, 等距采样是产生周期性现象的主要原因. 为了破坏等距采样, Kirchner 等提出了两种攻击方式.

(1) 基于几何失真和边界调制的攻击 (简称攻击 1). 该攻击在重采样点位置上添加 Gauss 扰动, 形成了新采样点 $(x'+\epsilon_x, y'+\epsilon_y)$, 其中 (ϵ_x, ϵ_y) 为零均值二维正态分布, 即 $f(\epsilon_x, \epsilon_y) = \frac{1}{2\pi\sigma_1\sigma_2}\mathrm{e}^{\frac{(\epsilon_x)^2}{(\sigma_1)^2}+\frac{(\epsilon_y)^2}{(\sigma_2)^2}}$. 标准差 σ_1, σ_2 越大, 攻击强度越大. 文献 [15] 中 $\sigma_1 = \sigma_2$, 下文利用 σ 表示攻击强度. 不等距的新采样点擦除了插值权值序列的周期性, 如图 1 第 3 行所示. 单独的几何攻击造成了反取证图像视觉质量的下降, 特别是在直线或边缘处留下了清晰可见的抖动噪声. 为了提高图像的视觉质量, 边界调制利用图像内容自适应地调整攻击强度. 具体来说, 边界或内容变化大的区域采用较小的攻击强度, 平滑区域采用较强的攻击强度. 当所有新采样点的位置确定后, 攻击 1 利用插值算法生成反取证图像. 从图 2 可以看出, 由攻击 1 生成的反取证图像 (图 2 第 1 行第 3 列) 与上采样图像 (图 2 第 1 行第 2 列) 在视觉上几乎没有差异, 但是反取证图像没有上采样的周期性痕迹. 图 2 显示反取证图像的 DFT 幅度谱 (第 2 行第 3 列) 和原始图像的 DFT 幅度谱 (第 2 行第 1 列) 几乎相同, 都没有显示周期性的亮点. 因此攻击 1 使得周期性的检测算法将反取证图像误判为原始图像.

(2) 双通道攻击 (简称攻击 2). 该攻击同时修改图像的低频和高频内容. 针对低频部分的攻击利用 5×5 中值滤波破坏重采样图像相邻像素间的线性关系. 针对高频部分的攻击则利用攻击 1 破坏相邻像素间的周期性. 首先, 提取原始图像与其 5×5 中值滤波图像的残差图像, 然后利用攻击 1 在残差图像上生成重采样图像的高频内容. 最后的反取证图像由攻击后的低频内容和高频内容相加形成.

攻击 1 和 2 都兼顾了攻击效果和反取证图像的视觉质量. 它们是目前针对重采样取证最好的反取证技术. 以往的重采样取证工作 [3~7,9,10] 都只关注取证特征的有效性, 而未考虑取证技术对抗恶意攻击的安全性. 下一节将提出一种安全有效的重采样取证特征.

3 重采样操作的特征提取

本文提出的安全重采样取证技术不仅能辨识重采样图像, 也能辨识经过反取证操作 [20] 的图像. 安全的重采样取证可以归纳为如下的假设检验:

H_0: 受测图像不是重采样图像;

H_1: 受测图像为重采样图像或经过反取证操作的图像.

H_0 类图像与 H_1 类图像间的差别主要体现在相邻像素间的不同依赖关系. 另外, H_1 类图像中的上采样图像和旋转图像也比 H_0 类图像更平滑.

H_1 类图像中的重采样图像由原始图像插值生成, 因此遗留下了插值操作的痕迹, 比如锯齿现象等. 由式 (2) 可知, 重采样点的像素值由其邻近的 l 个原始像素插值形成, l 的取值等同于插值核的宽度. 因此, 重采样图像中的相邻 l 个像素含有重叠的插值区域. 具体来说, 间隔为 i (i=1, 2, 3, ..., l−1) 的两个重采样像素拥有重叠的插值区域, 该区域含有 $l-i$ 个相同的原始图像素. 这些重叠区域导致了原始图像和重采样图像在相邻像素间的不同依赖关系.

H_1 类图像中的反取证图像也与原始图像呈现出不同的统计特征. 攻击 1 首先进行非等距采样, 然后插值形成反取证图像. 非等距采样会使某些区域聚集较多的采样点, 如图 1 第 3 行所示前两个重采样点的间隔较小, 而其余点的间隔较大. 攻击强度越大, 重采样点聚集的现象越严重. 聚集区域内的采

样点都是由几乎相同的原始像素插值生成, 它们也会呈现出局部模糊痕迹. 攻击 2 形成的反取证图像的高频内容由攻击 1 生成, 这会遗留下与攻击 1 类似的痕迹. 低频内容为经过 5×5 中值滤波的重采样图像. 中值滤波是一种非线性统计滤波器, 它不仅模糊图像, 也改变了图像相邻像素间的关系 [11,12]. 由此可推断由攻击 2 生成的反取证图像与原始图像在相邻像素间具有不同的依赖关系.

本文从重采样图像和反取证图像的高频成分提取检测特征. 高频成分含有的图像内容较少, 减少了自然图像内容多样性的影响. 此外, 高频成分也能反映重采样和反取证操作的模糊特性. 高频成分由图像与差分滤波器卷积生成. 灰度插值导致重采样图像在纹理内容上的统计特性发生了较大的变化. 自然图像的纹理往往具有多方向性, 因此本文从多方向的差分图像上提取特征.

3.1 多方向差分滤波器组

本文的多方向差分由图像与二阶差分滤波器 (图 3) 卷积生成. 使用更多的差分滤波器, 如一阶差分滤波器或高阶差分滤波器, 对取证算法性能的提升非常有限, 但是会造成特征提取的时间开销增加. 因此, 本文仅考虑二阶差分.

除了常见的水平方向 (horizontal, H)、垂直方向 (vertical, V)、对角线方向 (diagonal, D) 以及反对角线方向 (anti-diagonal, AD) 滤波器之外, 本文还考虑由上述 4 种方向的组合形成的滤波器. 假设图像的水平方向和垂直方向统计特性相似, 以及图像的对角线方向和反对角线方向统计特性相似, 可以把多方向差分滤波器分为 3 个滤波器组: (1) H, V 和 H+V 型滤波器组主要用于反映重采样操作在水平和垂直方向引起的变化, 用 P(1) 表示; (2) D, AD, D+AD 型滤波器组主要用于反映重采样操作在对角线和反对角线方向引起的变化, 用 P(2) 表示; (3) H+D, H+AD, V+D, V+AD 型滤波器主要用于反映重采样操作在水平 (垂直) 和对角线 (反对角线) 组合方向引起的变化, 用 P(3) 表示. 各组内的滤波器相互对称或旋转对称, 即组内各个滤波器可以由同组的任一滤波器通过对称或旋转生成. 最后共生成了 3 组滤波器 P(1)~P(3), 合计 28 种, 详见图 3.

本文从多方向差分提取的特征 (记为 F_{D}) 由直方图特征 (F_{H}) 和自回归系数 (F_{AR}) 组成. 直方图特征表征高频通道上的统计特征. 自回归系数用于揭示重采样图像以及反取证图像在相邻像素间的不同局部依赖关系.

$$g(t) = \sum_{i=1}^{p} \varphi(i) g(t-i) + \epsilon(t), \quad p < t < w, \tag{5}$$

其中 $g(t)$ 表示长度为 w 的随机序列, p 为拟合阶数, $\varphi(i)$ 表示自回归系数, $\epsilon(t)$ 为随机拟合误差.

3.2 特征提取

从各种差分图像提取自回归系数的方式由生成该图像的差分滤波器决定, 具体的提取方法如下: (1) 对于由图 3 的 H 滤波器、H+D 滤波器以及 H+AD 滤波器生成的差分图像, 首先将该图像按逐行扫描方式转换成一维行向量, 然后再估计该行向量的 AR 系数特征 F_{AR}; (2) 由图 3 中的 V 滤波器、V+D 滤波器以及 V+A 滤波器生成的差分图像, 首先将此图像按逐列扫描方式转换成一维列向量, 然后再估计该列向量的 AR 系数特征 F_{AR}; (3) 对于由其他滤波器生成的差分图像, 则按上述第 (1) 和 (2) 种方式分别估计此图像在行方向和列方向上的 AR 系数, 然后再将它们的平均值作为特征 F_{AR}.

考虑到直方图的对称性 (如图 4 所示), 直方图特征 F_{H} 的提取方式按式 (6) 提取:

$$F_{\mathrm{H}} = \left[h_0, \frac{h_{-1}+h_1}{2}, \ldots, \frac{h_{-q+1}+h_{q-1}}{2}, \frac{h_{-q}+h_q}{2} \right], \tag{6}$$

$$\text{P(1)[H]:}\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix};\qquad \text{P(1)[V]:}\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix};$$

$$\text{P(1)[H+V]:}\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix};$$

$$\text{P(2)[D]:}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix};\qquad \text{P(2)[AD]:}\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix};$$

$$\text{P(2)[D+AD]:}\ \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 0 & 1 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\text{P(3)[H+D]:}\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{pmatrix};\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix};$$

$$\text{P(3)[V+D]:}\ \begin{pmatrix} 1 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix};\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix};$$

$$\text{P(3)[H+AD]:}\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \\ 1 & 0 & 0 \end{pmatrix};$$

$$\text{P(3)[V+AD]:}\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix};\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

图 3 多方向滤波器组. 中括号内标示了滤波器的方向, 其中 H: 水平方向; V: 垂直方向; D: 对角线方向; AD: 反对角线方向; "+" 代表两个方向的组合

Figure 3 Multi-directional filters. H: horizontal; V: vertical; D: diagonal; AD: anti-diagonal; "+": union of two directions

这里 $h_{-q+1}, \ldots, h_0, h_1, \ldots, h_{q-1}$ 表示图像差分值的归一化频率, 而 h_{-q} 和 h_q 分别表示差分值小于等于 $-q$ 和大于等于 q 的归一化频率.

从 P(1)~P(3) 滤波器组生成的差分图像提取的特征用 $F_{\mathrm{D}}(i)$ $(i = 1, 2, 3)$ 表示, 由 $F_{\mathrm{AR}}(i)$ 以及 $F_{\mathrm{H}}(i)$ 组成. $F_{\mathrm{AR}}(i)$ 和 $F_{\mathrm{H}}(i)$ 分别为从第 i 组所有差分图像提取的 F_{AR} 和 F_{H} 的平均值. 由于组内滤

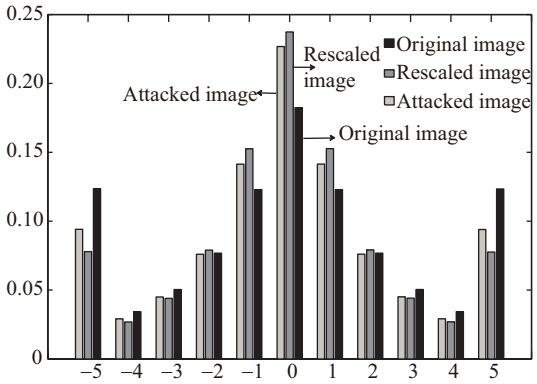

图 4 水平方向差分的统计直方图. 横坐标表示差分值, 纵坐标为归一化频率的均值. Original image: 原始图像; Rescaled image: 重采样图像; Attacked image: 反取证图像

Figure 4 The histogram of horizontal difference of image. The x axis is difference value. The y axis is average normalized frequency

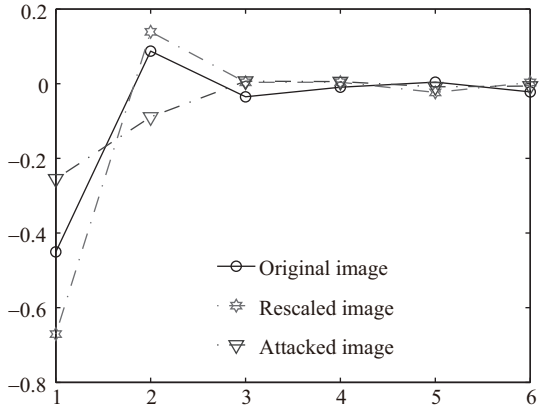

图 5 滤波器组 P(2) 的 6 阶 AR 系数特征. 横坐标表示 AR 系数编号, 纵坐标表示相应 AR 系数的均值

Figure 5 The 6-order AR coefficients of P(2). The x axis is index of AR coefficient. The y axis is average value of the AR coefficient

波器相互对称或旋转对称, 这种平均的特征表征方式既能减少特征元素间的冗余, 又极大地降低了特征的维数. 最后, 取证特征 $F_{\rm D}$ 由 P(1)~P(3) 提取的特征子集联合构成. $F_{\rm AR}(i)$ 和 $F_{\rm H}(i)$ 的维数分别是 p 和 $q+1$, 因此特征 $F_{\rm D}$ 的维数为 $3(p+q+1)$.

$$F_{\rm D}(i)=[F_{\rm AR}(i),F_{\rm H}(i)](i=1,2,3),F_{\rm D}=[F_{\rm D}(1),F_{\rm D}(2),F_{\rm D}(3)]. \tag{7}$$

直方图特征 $F_{\rm H}$ 的维数 q 从 BOSSRAW [1)]估计. 图 4 显示了各类图像在水平方向差分上的统计直方图. 直方图由 1500 幅灰度图像估计, 其中重采样图像为上采样图像, 尺度因子均匀地选择 [1.2,1.4,1.6,1.8], 插值算法也是均匀地选择双线性插值, 双三次插值和 Lanczos3 插值. 反取证图像采用的尺度因子和插值算法与重采样图像相同. 反取证图像由攻击 1 生成, 攻击强度 $\sigma=0.4$. 由图 4 可知绝大多数差分元素的取值位于 $[-5,5]$, 因而设置参数 $q=5$. 图 4 也反映了插值操作的模糊痕迹, 可以看出, 重采样图像和反取证图像在差分等于 −1, 0, 1 处的频率明显大于原始图像.

$F_{\rm AR}$ 反映了相邻像素间的局部依赖关系, 因此 AR 模型的拟合阶数 p 由插值核的宽度 l 确定. 常见的最近邻插值、双线性插值、双三次插值 (Bicubic) 以及 Lanczos3 插值核在一维方向的宽度分别为 1, 2, 4, 6. 为了覆盖尽可能多的插值算法, 本文设定 $p=6$. 图 5 显示了原始图像, 重采样图像以及反取证图像的 $F_{\rm AR}$ 在 BOSSRAW 图像库上的统计分布. 可以看出, 前 3 个 AR 系数明显地将原始图像和重采样图像以及反取证图像分开. 第 4, 5, 6 个 AR 系数的区分能力逐渐减弱. 可以预见的是第 6 个以后的 AR 系数的区分能力会更弱. 因此设定 $p=6$ 是合理的选择.

根据这些参数 p 和 q 的设置, 特征 $F_{\rm D}$ 的维数是 36. 本文所提出的检测算法的步骤如下: (1) 首先利用图 3 中的差分滤波器组与图像卷积生成多方向差分图像; (2) 在每个差分图像上提取 $F_{\rm AR}$ 和 $F_{\rm H}$; (3) 根据差分滤波器分组情况, 将各组内提取的特征值平均, 然后按式 (7) 组合成总特征 $F_{\rm D}$; (4) 分类器训练: 将训练图像提取的特征输入 SVM 训练以获取分类器; (5) 测试: 按步骤 (1)~(3) 提取测试图

1) "Break our steganographic system." http://boss.gipsa-lab.grenoble-inp.fr/.

像的总特征 F_{D}, 然后输入分类器检测图像是否经过重采样操作或反取证操作.

4 实验结果

本文所用的混合图像库来自两个常用的图像库: (1) BOSSRAW 图像库. 该图像库包含 10000 张 raw 格式的彩色图片, 这些图片由佳能、宾得、尼康等相机拍摄. 该图像库在文献 [9] 中被用作测试数据库. (2) DID 图像库 [22]. 此图像库包含从多部相机拍摄的一万多张大尺寸 JPEG 格式图像, 常用于相机源取证和数字图像操作取证 [10,12]. 它们所包含的图像都未经过重采样操作. 每个图像库随机地选择 1500 幅图像并抽取灰度分量. 最后得到了包含 3000 幅灰度图像的混合图像库.

本文的实验考虑了两种常见的重采样操作: 尺度变换 (包括放大和缩小) 和旋转. 原始图像、重采样图像以及反取证图像的大小皆为 512×512. 它们都是通过截取相应图像的中间部分生成. 原始图像库和重采样图像库或反取证图像库形成了一个训练测试组. 对于每一种训练测试组, 随机选择一半的图像用于 SVM 分类器训练, 余下的一半图像用于测试.

Gauss 核的 SVM [23] 被用作分类器, 为了避免训练中的过拟合现象, 采用五重校验法寻找最佳的参数 C (惩罚参数) 和 γ (Gauss 核参数). 实验结果的评估采用受试者工作特征曲线 ROC 和最小平均检测错误率 P_e [11,12].

$$P_e=\min\left(\frac{1-\mathrm{TPR}+\mathrm{FPR}}{2}\right), \tag{8}$$

其中 FPR 为虚警率 (即将 H_0 类图像误认为 H_1 类图像的比率), TPR 为正确阳性检测率 (即将 H_1 类图像识别为 H_1 类图像的比率). 在测试样本正负样例相等的情况下, 最小平均检测错误率 P_e (以下简称检测错误率) 越小, 检测器性能越好.

Li 等 [10] 在频率域提出的矩特征 (moment feature, 简记为 F_{M}) 是目前已知性能较好的重采样取证特征. 本文将从有效性、安全性以及对抗 JPEG 压缩的鲁棒性方面与 F_{M} 特征对比. 为了保证对比的公平性, 所有算法的实验设置都相同.

4.1 特征有效性评估

本小节实验测试所提方法是否能够区分原始图像和重采样图像. 重采样操作的参数包括尺度因子 (旋转角度) 和插值算法. 已有的重采样取证算法 [3~7,9,10] 表明尺度因子是影响取证效果的关键参数, 其中取证算法 [3~7] 等在下采样图像上的取证效果都不好. 本小节实验考虑了下采样、上采样和旋转 3 种图像库. 每种图像库拥有 3000 张重采样图像. 每种图像库包含 12 类数量相等的图像, 对应着 4 种尺度因子或旋转角度, 3 种常见的插值算法: 双线性插值、双三次插值以及 Lanczos3 插值. 下采样、上采样和旋转操作的因子分别为 [0.6,0.7,0.8,0.9], [1.2,1.4,1.6,1.8] 和 $[\frac{\pi}{36}, \frac{\pi}{18}, \frac{\pi}{12}, \frac{\pi}{9}]$.

该实验使用了 3 个训练测试组, 分别测试所提方法 F_{D} 和对比方法 F_{M} 在无 JPEG 压缩下区分 (1) 原始图像和上采样图像; (2) 原始图像和下采样图像; (3) 原始图像和旋转图像的能力. 图 6 中的 ROC 曲线显示所提方法能有效地区分原始图像与重采样图像, 其中在上采样的测验中取得了最小的检测错误率 ($P_e=0.6\%$). 在实际应用中, 检测器在低虚警率下的检测准确率更令人关注. 如图 6 所示, 在 $\mathrm{FPR}=1\%$ 的情况下, 所提特征在检测上采样、下采样和旋转操作取得的 TPR 分别为 99.7%, 98.1%, 93.5%, 分别超出 F_{M} 特征 27.1%, 53.7%, 58.4%. 这些实验结果表明, 本文所提特征大幅度地提高了重采样检测器在下采样操作和旋转操作中的检测性能.

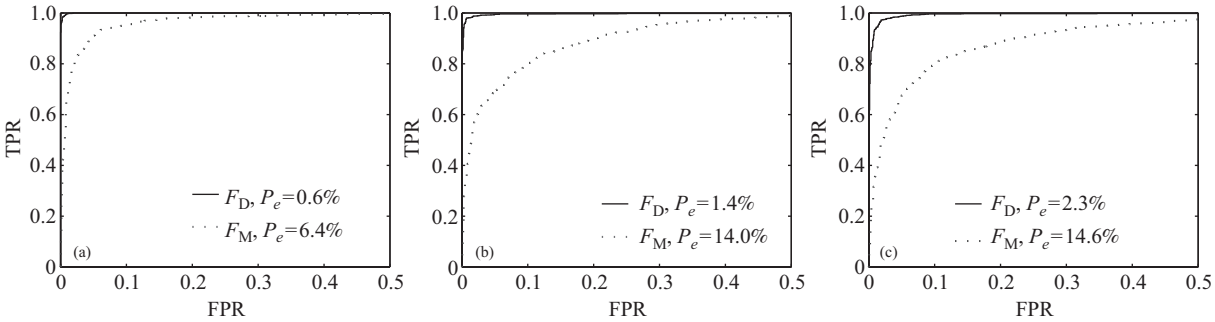

图 6 特征 F_D 和 F_M 的 ROC 曲线, 其中 (a)∼(c) 分别检测上采样、下采样以及旋转图像

Figure 6 The ROC of F_D and F_M for (a) up-scaling (b) down-scaling and (c) rotation

表 1 评估 F_D 和 F_M 区分原始图像与重采样图像的性能 P_e (%)

Table 1 Estimation the ability of F_D 和 F_M of classifying original image and resampling image via P_e (%)

Feature	Attack type																	
	Attack 1									Attack 2								
	Up-scaling			Down-scaling			Rotation			Up-scaling			Down-scaling			Rotation		
	$\sigma = 0.3$	0.4	0.5	0.3	0.4	0.5	0.3	0.4	0.5	0.3	0.4	0.5	0.3	0.4	0.5	0.3	0.4	0.5
F_D	0.03	0	0	0.3	0.07	0.07	0.03	0	0	0.2	0.03	0.03	1.1	0.5	0.4	0.8	0.3	0.3
F_M	3.6	2.5	2.3	10.9	8.5	6.4	0.8	5.2	3.5	6.3	4.0	2.5	12.6	10.5	8.4	12.9	8.0	5.9

4.2 安全性能评估

本小节实验测试所提特征 F_D 能否区分原始图像与反取证图像. 攻击强度 σ 在反取证操作中起着平衡攻击效果和被攻击图像质量的作用. 聪明的攻击者会根据图像内容自适应地调节. 比如, 对于平滑区域采用较大的 σ, 而纹理或边界则采用较小的 σ. 为此, 除了反取证算法 [20] 使用的默认攻击强度 $\sigma = 0.4$, 本文还考虑相对较弱 ($\sigma = 0.3$) 与较强攻击强度 ($\sigma = 0.5$). 反取证操作的其他参数包括两种攻击 (攻击 1 和攻击 2) 和 3 种重采样操作 (下采样、上采样和旋转), 共计生成 18 种图像库. 每种反取证图像库拥有 3000 张图像, 共计生成了 54000 张反取证图像. 尺度因子 (旋转角度) 与插值算法采用了与 4.1 小节相同的设置.

表 1 详细地列出了所提特征 F_D 在 18 个训练测试图像组上的检测结果. 每个训练测试组由 3000 张原始图像和 3000 张反取证图像组成. 如表 1 所示, 基于特征 F_D 的检测器能有效地区分原始图像与反取证图像. 在所有测试组中, F_D 的检测错误率 P_e 位于区间 [0, 1.1%], 在其中 4 个测试组上取得了完美的测试效果 ($P_e = 0\%$). 实验结果表明 P_e 与攻击强度 σ 成反比. 攻击强度越大, 反取证操作篡改的内容也越多, 因此越容易被检测. 总体来说, 基于特征 F_D 的检测器在各种反取证操作的检测中取得了有效且稳定的检测效果. F_M 的检测结果表明 F_M 也能区分原始图像与反取证图像, 但是其检测效果明显低于 F_D. 对比 4.1 和 4.2 小节的实验结果可知, 反取证图像比重采样图像更容易被检测. 这是因为反取证图像既遗留了重采样图像的部分痕迹, 同时反取证操作又带来了新的痕迹, 这加剧了原始图像与反取证图像之间的差异.

表 2 评估 F_{D} 和 F_{M} 对抗 JPEG 压缩的鲁棒性能 P_e (%)

Table 2 Estimation the robustness of F_{D} 和 F_{M} against JPEG compression P_e (%)

JPEG quality factor	Classifiers	Resampling type	F_{D}	F_{M}
JPEG 95	Original vs Rescaled	Up-scaling	3.0	10.3
		Down-scaling	7.2	18.5
		Rotation	7.2	21.5
	Original vs Attacked	Up-scaling	1.4	7.7
		Down-scaling	3.8	14.8
		Rotation	2.4	13.6
JPEG 90	Original vs Rescaled	Up-scaling	7.3	14.1
		Down-scaling	14.0	23.3
		Rotation	12.3	25.7
	Original vs Attacked	Up-scaling	3.3	12.6
		Down-scaling	9.2	20.3
		Rotation	5.1	18.3

4.3 对抗 JPEG 压缩的鲁棒性能评估

本小节实验测试 F_{D} 对抗 JPEG 压缩的鲁棒性能. JPEG 压缩是常用的图像保存格式. 有损 JPEG 压缩不可避免地擦除了图像操作的部分痕迹. 一般来说, 压缩质量因子越低, 抹除的痕迹越多. 已有的重采样取证算法 [3~6] 对抗 JPEG 压缩的鲁棒性能都不好. 本小节考虑了 JPEG 95 和 JPEG 90 两种质量因子. JPEG 压缩为后处理操作, 即首先对图像进行重采样或反取证操作, 然后再 JPEG 压缩.

首先测试 F_{D} 在 JPEG 压缩下区分原始图像与重采样图像的能力. 训练测试组中的原始图像和重采样图像都经过了 JPEG 压缩, 其中重采样图像库采用了 4.1 小节实验所采用的 3 个重采样图像库. 表 2 的检测结果表明基于 F_{D} 的检测器对抗 JPEG 压缩的鲁棒性能较好. 所提特征在 JPEG 95 压缩下的上采样测试中取得了最小的检测错误率 $P_e = 3.0\%$, 其检测错误率远低于 F_{M}. 在所有 6 个测试训练组中, F_{D} 的检测错误率至少比 F_{M} 低 6.8%.

紧接着测试 F_{D} 在 JPEG 压缩下区分原始图像与反取证图像的能力. 表 1 的实验结果表明尺度因子是影响检测效果的关键因素. 因此实验考虑了上采样、下采样和旋转 3 种反取证图像库. 每个图像都包含 18 类数量相等的反取证图像, 分别对应着两种攻击 (攻击 1 和攻击 2)、3 种攻击强度 ($\sigma = 0.3, 0.4, 0.5$) 和 3 种插值算法 (双线性插值、双三次插值、Lanczos3 插值). 所有训练测试组中的图像都经过了 JPEG 95 和 JPEG 90 压缩. 表 2 的实验结果显示基于 F_{D} 的检测器能有效地对抗 JPEG 压缩. 在 JPEG 95 压缩下, F_{D} 在上采样、下采样以及旋转操作的检测错误率分别为 1.4%, 3.8% 和 2.4%, 比 F_{M} 分别低 6.3%, 11.0% 和 11.2%.

4.4 区分 H_0 类图像与 H_1 类图像

本节测试 F_{D} 区分 H_0 类图像 (原始图像) 与 H_1 类图像 (重采样图像和反取证图像) 的能力. 以往的重采样检测器只考虑了区分原始图像与重采样图像 [3~7,9,10]. 这种检测器往往会将反取证图像误判为原始图像. 由 H_0 类图像与 H_1 类图像作为训练图像生成的检测器能有效地解决反取证图像的误判问题. 但是这类检测器对取证特征提出了较高的要求. 图 7 显示了利用线性判别分类器 LDA 分别将 BOSSRAW 图像库中的原始图像、重采样图像和反取证图像的特征 F_{D} 降低到 2 维的散点图分布.

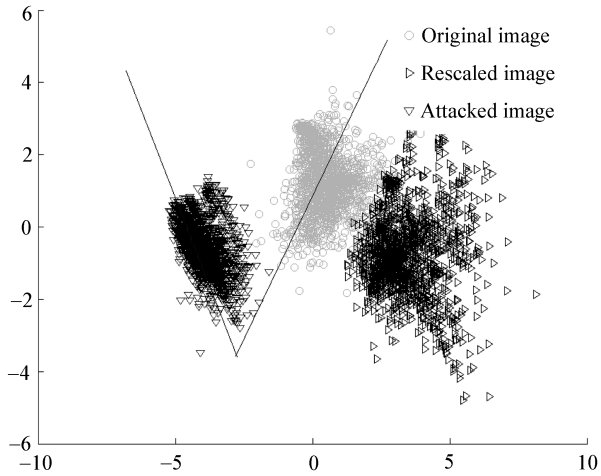

图 7　原始图像、重采样图像和反取证图像的二维散点图分布. 横纵坐标分别表示降维后得到的第 1 和第 2 维特征

Figure 7　The scatter of original image, rescaled image and attacked image.The x (y) is the first(second)element of the feature after dimension reduction

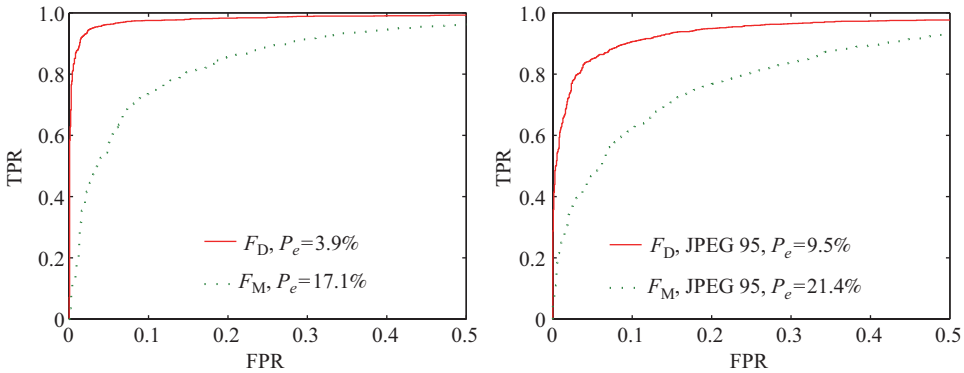

图 8　(网络版彩图) 区分 H_0 类图像与 H_1 类图像的 ROC 曲线. (a) 无 JPEG 压缩, (b) JPEG 95 压缩图像

Figure 8　(Color online) The ROC of classifying H_0 from H_1 under (a) without JPEG compression and (b) JPEG 95 compression

可以看出 H_0 类图像与 H_1 类图像之间有清晰的边界分割线, 这预示着 $F_{\rm D}$ 能区分 H_0 类图像与 H_1 类图像.

为了模拟实际的应用场景, 考虑了多样性的 H_1 类图像. H_1 类图像中的重采样图像来自 4.1 小节所采用的 3 个重采样图像库, 每个图像库选取了 500 副图像. H_1 类图像中的反取证图像由 4.1 小节的 18 种反取证图像组成, 每类反取证图像的数量大致相等. 该实验的训练测试组由 3000 张 H_0 类图像和 3000 张 H_1 类图像 (1500 张重采样图像 + 1500 张反取证图像) 组成. 同时考虑无 JPEG 压缩和 JPEG 95 两种情形. 随机地选择训练测试组内一半的 H_0 类图像和 H_1 类图像训练 SVM 分类器, 余下的一半用于测试.

图 8 显示基于 $F_{\rm D}$ 的分类器能有效地区分 H_0 类图像与 H_1 类图像. 如图 8(a) 所示, 在 FPR = 1% 的情况下, 所提特征取得的 TPR 为 87.9%, 超出 $F_{\rm M}$ 66.4%. 即使在 JPEG 95 压缩的情况下 (图 8(b)), 基于 $F_{\rm D}$ 的分类器也取得了较低的检测错误率 $P_e = 9.5\%$.

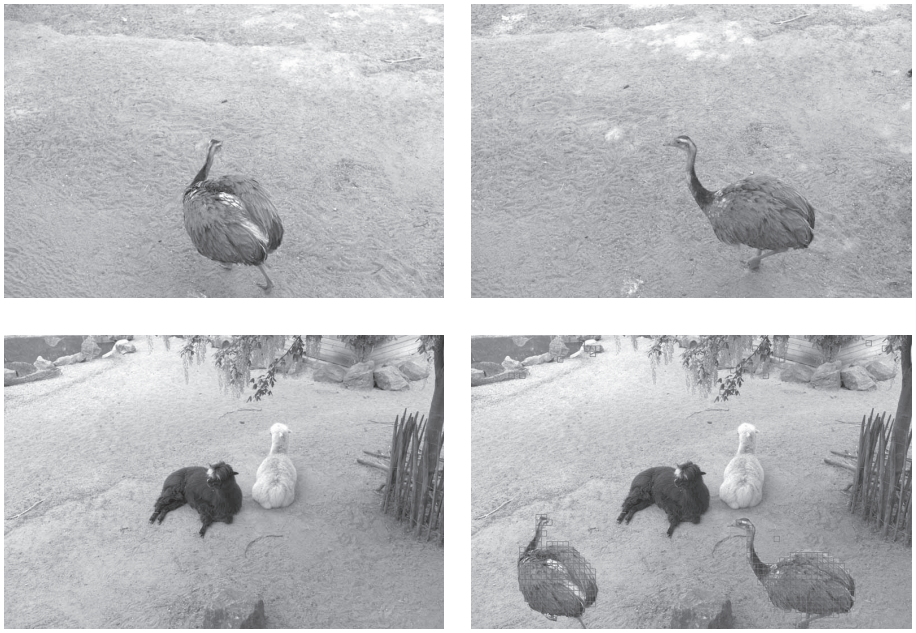

图 9 检测篡改图像. 前 3 个图像为原始图像. 第 4 个图像为篡改图像, 它由拼接第 1 和 2 个图像中的鸟到第 3 个图像形成. 篡改图像中位于左边和右边的鸟分别经过反取证和重采样操作. 检测出的篡改区域用方框标示

Figure 9 Example of tampering detection. The forged image (fourth) is spliced the bird in the first and second to the third image. The left bird in the forged image is undergone attack 2. The right bird in the forged image is undergone a series of rescaling. The detected tampered region is marked by box

4.5 篡改检测实例

本小节利用 F_{D} 进行篡改检测. 图 9 中的第 4 个图像是一个篡改实例. 它由拼接第 1 个和第 2 个图像中的鸟到第 3 个图像合成. 篡改图像中位于左边的鸟由反取证操作攻击 2 生成. 这些参数采用了文献 [20] 中的默认参数 (双线性, $\sigma = 0.4, s = 0.8$). 篡改图像中位于右边的鸟依次经过了尺度因子 $s = 0.7$ 的下采样和 $s = 1.2$ 的上采样. 两种采样方式都使用双三次插值算法. 图 9 的所有图像为 TIFF 图像, 尺寸为 5216×3472.

篡改检测的步骤依次为 (1) 利用特征 F_{D} 训练 SVM 分类器; (2) 将待测图像分块, 块大小为 64×64; (3) 提取各块的特征 F_{D}; (4) 利用分类器判别各个分块是否经过了篡改操作. SVM 分类器通过训练尺寸为 64×64 的图像块生成. 图像块由截取原图像的中心部分得到. 训练图像库由 H_0 类混合图像库 (3000 张原始图像) 和相应的 H_1 类图像库 (1500 张重采样图像 + 1500 张反取证图像) 构成. 重采样图像和反取证图像的参数与 4.4 小节的设置相同. 在实际应用中, 测试图像很有可能不在训练图像库中. 为了模拟实际应用场景, 训练的图像数据库不包含测试实例图 9 中的所有 4 个图像.

图 9 中的第 4 个图像显示了篡改检测的实验结果, 方框标示了篡改区域. 可以看出, F_{D} 较准确地标示了大多数的篡改区域. 在细长的部分, 比如颈部、嘴角、腿脚等出现了漏检的情况. 这是因为细长区域的分块同时含有篡改区域和未篡改的区域. 分类器对于上述分块的检测准确率较差. 图 9 的实验结果中也存在少量的虚警结果 (将未篡改的区域误判为篡改区域). 但是这些虚警区域零星孤立分布, 可以根据常识将这些虚警区域去掉.

图 9 的实例是一个简单的篡改实例. 聪明的篡改者会采用其他手段以规避取证算法的检测. 该实例的检测结果表明重采样取证算法能识别经过重采样的篡改. 这预示着综合类的取证算法 (即针对多

种图像操作的取证) 或许能检测更为复杂的篡改操作. 这些内容将是我们下一步的研究方向.

5 结论

本文提出了一种新的重采样取证特征, 第一次同时考虑了重采样取证的有效性和安全性问题. 所提特征由自回归模型的系数和直方图组成, 反映了重采样操作和反取证操作由于灰度插值而生成的痕迹. 我们从多方向差分图像提取特征, 并根据差分图像间的对称性对所提特征降维.

在多个图像数据库组成的混合图像库上的测试结果表明所提出的集成检测器既能有效地检测重采样操作, 也能检测已有的反取证操作. 所提检测器大幅度地提高了下采样操作和旋转操作的检测准确率, 改善了已有重采样取证算法对抗 JPEG 压缩的鲁棒性. 篡改图像上的检测结果进一步验证了所提检测器的有效性和安全性.

重采样是篡改图像中频繁使用的一种操作. 篡改者往往还会同时使用其他操作美化篡改图像. 分析其他操作 (JPEG 压缩、对比度增强、锐化、平滑滤波等) 与重采样操作混合后的痕迹将有利于检测篡改操作. 这些将会是我们的后续研究方向.

参考文献

1 Gong D F, Liu F L, Luo X Y. An image edge based robust watermarking algorithm. Sci Sin Inform, 2013, 43: 1410–1430 [巩道福, 刘粉林, 罗向阳. 一种基于图像边缘的鲁棒水印算法. 中国科学: 信息科学, 2013, 43: 1410–1430]

2 Xiang S J, Yang J Q, Huang J W. Perceptual video hashing robust against geometric distortions. Sci Sin Inform, 2012, 42: 578–587 [项世军, 杨建权, 黄继武. 抗几何失真的视频 Hashing 算法研究. 中国科学: 信息科学, 2012, 42: 578–587]

3 Popescu A C, Farid H. Exposing digital forgeries by detecting traces of resampling. IEEE Trans Signal Process, 2005, 53: 758–767

4 Gallagher A C. Detection of linear and cubic interpolation in JPEG compressed images. In: Proceedings of the 2nd Canadian Conference on Computer and Robot Vision, Canada, 2005. 65–72

5 Mahdian B, Saic S. Blind authentication using periodic properties of interpolation. IEEE Trans Inf Foren Sec, 2008, 3: 529–538

6 Kirchner M. Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. In: Proceedings of the 10th ACM Workshop on Multimedia and Security, Oxford, 2008. 11–20

7 Wei W M, Wang S Z, Zhang X P, et al. Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery. IEEE Trans Inf Foren Sec, 2010, 5: 507–517

8 Cao G, Zhao Y, Ni R R. Forensic identification of resampling operators: a semi non-intrusive approach. Foren Sci Int, 2012, 216: 29–36

9 Feng X Y, Cox Ingemar J, Doerr G. An energy-based method for the forensic detection of re-sampled images. In: Proceedings of the 12th International Conference on Multimedia and Expo, Barcelona, 2011. 1–6

10 Li L, Xue J R, Tian Z Q, et al. Moment feature based forensic detection of resampled digital images. In: Proceedings of the ACM Multimedia Conference, Barcelona, 2013. 569–572

11 Yuan H D. Blind forensics of median filtering in digital images. IEEE Trans Inf Foren Sec, 2011, 6: 1335–1345

12 Chen C L, Ni J Q, Huang R B, et al. Blind median filtering detection using statistics in difference domain. In: Information Hiding. Berlin: Springer, 2012. 1–15

13 Stamm M, Liu K J R. Forensic estimation and reconstruction of a contrast enhancement mapping. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, 2010. 1698–1701

14 Cao G, Zhao Y, Ni R R, et al. Unsharp masking sharpening detection via overshoot artifacts analysis. IEEE Signal Process Lett, 2011, 18: 603–606

15 Luo W Q, Huang J W, Qiu G P. JPEG error analysis and its applications to digital image forensics. IEEE Trans Inf Foren Sec, 2010, 5: 480–491

16 Jiang X H, Wang W, Sun T F, et al. Detection of double compression in MPEG-4 videos based on markov statistics. IEEE Signal Process Lett, 2013, 20: 447–450
17 Amerini I, Ballan L. A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans Inf Foren Sec, 2011, 6: 1099–1110
18 Yang R, Luo W Q, Huang J W. Multimedia forensics. Sci Sin Inform, 2013, 43: 1654–1672 [杨锐, 骆伟祺, 黄继武. 多媒体取证. 中国科学: 信息科学, 2013, 43: 1654–1672]
19 Kodovsky J, Fridrich J. Steganalysis in resized images. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, 2013. 2857–2861
20 Kirchner M, Bohme R. Hiding traces of resampling in digital images. IEEE Trans Inf Foren Sec, 2008, 3: 582–592
21 Bohme R, Kirchner M. Counter-forensics: attacking image forensics. In: Digital Image Forensics. New York: Springer, 2013. 327–366
22 Gole T, Bohme R. Dresden image database for benchmarking digital image forensics. In: Proceedings of the ACM Symposium on Applied Computing, Sierre, 2010. 1584–1590
23 Chang C, Lin C. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Tech, 2011, 3: 12–27

Resampling forensics based on multi-directional difference

Anjie PENG[1,2 †], Hui ZENG[1,3 †] & Xiangui KANG[1,3*]

1 *School of Data and Computer Science, Sun Yat-Sen University, Guangzhou* 510006*, China*;
2 *School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang* 621010*, China*;
3 *Guangdong Key Laboratory of Information Security, Guangzhou* 510006*, China*
*E-mail: isskxg@mail.sysu.edu.cn
† These authors contributed equally.

Abstract The forensics of resampling has recently drawn much attention since it has been used in image tampering to beautify the forged image. Existing detectors work well without malicious adversary. However, this is not possible in real life. The attack targeted against the periodicity-based method has successfully defeated the forensics of resampling. In this paper, we propose a novel detection of resampling that extracts the feature from the image difference domain. To achieve an effective and secure detector of resampling, multi-directional differences are calculated and divided into three groups. We fit each multi-directional difference into an auto-regressive (AR) model. The AR parameters and histograms are averaged within each group and are concatenated to form the final feature. Experimental results on a large composite image database verify that the proposed detector is not only effective in resampling detection, but also can resist the well-known attacks. The proposed method performs much better than state-of-the-art methods, achieving significant improvement in the detection of down-sampling image and rotated image. Furthermore, the proposed method greatly enhances the detector's robustness against JPEG compression. The detection of tampering again verifies the effectiveness and security of the proposed feature.

Keywords digital image forensics, resampling, multi-directional difference, auto-regressive model

Anjie PENG was born in 1981. He received his B.S. and M.S. degrees from Sun Yat-sen University in 2005 and 2007, respectively. Currently, he is a Ph.D. candidate at Sun Yat-sen University. His research interests include multimedia forensics and security, and machine learning.

Xiangui KANG received his B.S. degree from Peking University, Beijing, China, in 1990, M.S. degree from Nanjing University, Nanjing, China in 1993, and Ph.D. degree from Sun Yat-Sen University, Guangzhou, China in 2004. He is currently a professor at the School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China. His research interests include information forensics, watermarking, and multimedia communications and security.

Hui ZENG was born in 1982. He received his master's degree from Nanjing University of Posts and Telecommunications, Nanjing in 2007. Currently, he is a Ph.D. candidate at Sun Yat-sen University. His research interests include multimedia forensics and game theory.