文章编号:1001 - 9081(2009)06 - 1617 - 05

适用于多区域篡改的 JPEG 图像认证算法

高宝建,冯 峰,侯爱琴
(西北大学 信息科学与技术学院,西安 710127)
(fengfeng2536@163.com)

摘 要:利用纠错编码的陪集分解原理,给出了一种从长度为 $2^m - 1$ 的二进制信息序列中提取 m 比特摘要的方法,证明了 JPEG 压缩引起图像块均值失真的一个上界,并据此提出一种新的图像预处理方法。在此基础上,设计了一种新的基于数字签名的近似图像认证算法。在 JPEG 质量因子大于等于 25 的情况下,该算法可以区分篡改和正常压缩失真;在多区域篡改的情况下,该算法可以实现篡改检测和准确定位,且具有较高的篡改检测概率和较低的虚警概率。理论分析和仿真实验均证明了该算法的正确性和有效性。

关键词:图像认证;篡改检测与定位;JPEG 压缩
中图分类号: TP391 **文献标志码**:A

New JPEG image authentication algorithm applied to multi-region tampering

GAO Bao-jian, FENG Feng, HOU Ai-qin
(*School of Information Science and Technology, Northwest University, Xi'an Shaanxi* 710127, *China*)

Abstract: A new method of distilling m bits abstract from the binary information sequence with the length of $2^m - 1$ was proposed based on the coset decomposing theory of correct error coding. An important conclusion was proved that distortion of image block mean caused by JPEG lossy compression has a maximal value; a new image pre-processing method based on the conclusion was introduced and analyzed. Based on those, a new image authentication algorithm based on digital signature was proposed. It can distinguish between JPEG lossy compression and malicious tampering when the JPEG quality factor is greater than or equal to 25. When more than one region was tampered, it not only can effectively detect and localize the tampered area, but also has high probability of tampering detection and lower probability of false alarm. The results of theoretical analysis and emulation show that this algorithm is correct and effective.

Key words: image authentication; tamper detection and location; JPEG compression

0 引言

一般来说,图像认证有两种方法。一种是将图像转换为二进制码流,利用传统的数据认证方法[1]来实现认证,这种方法对数据流的改变高度敏感,只要数据流有一位发生改变,该认证方法都会认为图像被篡改而拒绝接收,所以我们将这种方法称为严格(strict)图像认证;另一种方法是基于图像内容的完整性和人类视觉的迟钝性的方法,这种方法对图像的正常处理(例如压缩和滤波等)不敏感,所以我们将这种方法称为非严格(nonstrict)或者近似(approximate)图像认证。近似图像认证不仅要求能区分正常图像处理和篡改,而且要能对篡改进行准确定位[2]。不难看出近似图像认证方法更具有实际意义。

目前,近似图像认证技术已取得了较大的发展,提出了很多方案。分析这些方案,主要包括基于半脆弱数字水印的方法和基于数字签名的方法。文献[3-6]属于基于半脆弱数字水印的近似图像认证方案,虽然没有签名或者摘要的秘密传输问题,但是又要面对各种水印攻击,一般情况下对图像有损压缩的鲁棒性不是很强,只能抵抗 JPEG 质量因子(Quality Factor, QF)≥50 的压缩。文献[2,7]利用图像块均值提出了一种基于数字签名的近似图像认证方案,该方案能较好地区分篡改与压缩失真,但其篡改检测概率相对较低,只有 25%,只能抵抗 $QF \geqslant 50$ 的压缩。分析发现,无论是基于数字签名的方法还是基于半脆弱数字水印的方法,目前的研究方案大都是针对一个局部篡改区域的情况而进行的,当篡改区域多于一个时,这些方案都无法正常工作,而对图像的多区域篡改在现实生活中是非常普遍的,例如改变图像中某重要部分的位置等,所以有必要研究图像多区域篡改的认证问题。

本文针对已有认证方案存在的不足,提出了一种新的基于数字签名的近似图像认证算法。在该算法中,设计了新的摘要提取方法和图像预处理方法,针对图像多区域篡改的情况,设计了一种区域检测定位和交叉检测定位复合的篡改检测定位方法,从而使本文算法在 $QF \geqslant 25$ 的情况下,可以区分篡改和压缩失真,在多区域篡改的情况下,可以实现篡改检测和准确定位,且具有较高的篡改检测概率和较低的虚警概率。

1 认证摘要提取的新方法

在图像内容认证中,虽然通过 hash 函数提取摘要可以提高安全性,但是这种方法提取的摘要长度为 128 比特或者 160 比特,在图像内容认证中直接使用,会使总的认证摘要很长,一般要从 128 比特或者 160 比特中再次提取较短的摘要,这样可大幅减少摘要长度,但是付出的代价是篡改检测概率

收稿日期:2008 - 12 - 11;**修回日期**:2009 - 02 - 25。 **基金项目**:陕西省国际科学与技术合作项目(2006KW-21)。

作者简介:高宝建(1963 -),男,陕西宝鸡人,副教授,硕士,主要研究方向:信号与信息处理、图像认证、信息隐藏; 冯峰(1983 -),男,陕西榆林人,硕士研究生,主要研究方向:信号与信息处理、图像认证; 侯爱琴(1967 -),女,陕西渭南人,讲师,硕士,主要研究方向:信号与信息处理、信息安全。

也大幅下降(分析见4.2节)。本文基于汉明码的代数结构,提出一种新的摘要提取方法,该方法在保证安全性及对原始信息较高敏感性的情况下,摘要长度为 lb $(n+1)$,n 为信息序列长度。

1.1 认证摘要提取算法

对任意给定的 $n(n=2^m-1)$ 位二进制特征数据 $\boldsymbol{a}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\cdots,\boldsymbol{a}_n)$,选定一列向量遍历 $n=2^m-1$(全零除外)种状态的 $m\times n$ 矩阵 $\boldsymbol{H}=(\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3,\cdots,\boldsymbol{h}_n)$,$\boldsymbol{h}_i$ 为 m 维列向量,设 $\boldsymbol{c}=(\boldsymbol{c}_1,\boldsymbol{c}_2,\boldsymbol{c}_3,\cdots,\boldsymbol{c}_m)$ 为一给定的 m 维行向量。

认证摘要提取算法如下:

第1步 计算 $\boldsymbol{s}^{\mathrm{T}}=\boldsymbol{H}\,\boldsymbol{a}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$,$\boldsymbol{s}=(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\cdots,\boldsymbol{s}_m)$ 为 m 维行向量;

第2步 计算 $\boldsymbol{p}_i^{\mathrm{T}}=\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{h}_i(i=1,2,\cdots,n)$,$\boldsymbol{p}_i$ 为 m 维行向量;

第3步 认证摘要提取:

```
Za = 0;
for i = 1:n
  if p_i^T = =0
    Za = i;
  end
end
```

则 Za 为和二进制特征数据 $\boldsymbol{a}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\cdots,\boldsymbol{a}_n)$ 对应的认证摘要,记为 $Za=\Gamma(\boldsymbol{a},\boldsymbol{c},\boldsymbol{H})$,且 Za 为集合$\{0,1,2,\cdots,2^m-1\}$中的元素,其二进制表示的长度为 m。

1.2 算法性能分析

这里我们主要分析本算法得到的认证摘要对特征数据变化的敏感性。

依据纠错码理论[8],由上述算法的 $\boldsymbol{H}$ 矩阵的组成,我们可以将其看作$(n,k)=(2^m-1,2^m-1-m)$汉明码的效验矩阵,由算法知,$\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}=\boldsymbol{H}\,\boldsymbol{a}^{\mathrm{T}}$,故 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 为 n 维二进制特征数据 $\boldsymbol{a}^{\mathrm{T}}$ 的伴随式;由于汉明码是完备码,所以可以把整个 n 维线性空间的 2^n 个元素按伴随式划分为 2^m 个互不重叠的陪集,每个陪集包含 2^k 元素,每个陪集对应一个伴随式;当伴随式 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 给定时,如果和其对应的 n 维二进制特征数据 $\boldsymbol{a}^{\mathrm{T}}$ 出错,则当出错的 $\boldsymbol{a}^{\mathrm{T}}$ 仍然在和给定的伴随式 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 对应的陪集内时,伴随式 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 不变,否则伴随式 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 一定改变;由于 $\boldsymbol{c}^{\mathrm{T}}$ 不变,所以伴随式 $\boldsymbol{s}^{\mathrm{T}}\oplus\boldsymbol{c}^{\mathrm{T}}$ 的变化必然对应 $\boldsymbol{s}^{\mathrm{T}}$ 的变化,由算法知,$\boldsymbol{s}^{\mathrm{T}}$ 的变化必然引起认证摘要 Za 的变化;这样我们不难得到 n 维二进制特征数据 $\boldsymbol{a}^{\mathrm{T}}$ 变化时,其认证摘要 Za 改变的概率为:

$$P=(2^m-1)\,2^k/2^n=(2^m-1)/2^{n-k}=(2^m-1)/2^m \tag{1}$$

例如:当 $m=7$ 时,$P=99.2\%$;当 $m=8$ 时,$P=99.6\%$。

由此,我们可以得出结论,本算法得到的认证摘要对特征数据的变化高度敏感,这正是认证摘要所必需的。同时也不难看出,相对 hash 函数提取摘要的方法[7],本算法实现简单灵活,摘要长度短;而 $\boldsymbol{c}^{\mathrm{T}}$ 的引入提高了算法的安全性,如果将 $\boldsymbol{c}^{\mathrm{T}}$ 和 $\boldsymbol{H}$ 矩阵作为密钥,其密钥空间的大小为 $2^m\times(2^m-1)!$,具有较高的安全性。

2 图像特征量提取及预处理方法

2.1 JPEG 压缩对图像块均值的影响

将原始灰度图像 I 分成 $w\times h$ 个互不重叠的块 $B_{i,j}(0<i<w-1,0<j<h-1)$,块大小为 8×8 像素。块内64像素灰度值的平均值称为块均值,记为 $\tilde{m}_{i,j}=\frac{1}{64}\sum_{k=1}^{8}\sum_{l=1}^{8}b_{k,l}^{i,j}$,$\tilde{m}_{i,j}$ 一般不为整数,对其四舍五入取整得 $m_{i,j}$,$m_{i,j}$ 称为整数型块均值;对 $B_{i,j}$ 作离散余弦变换(Discrete Cosine Transform, DCT),其直流系数为 DC_{ij},用量化阶 q 对 DC_{ij} 量化,$m'_{i,j}$ 表示量化后IDCT图像块均值四舍五入取整的结果,$m'_{i,j}$ 称为量化后整数型块均值。

定理1 将给定灰度图像 I 分成 $w\times h$ 个互不重叠的 8×8 块 $B_{i,j}(0<i<w-1,0<j<h-1)$,对 $B_{i,j}$ 作DCT变换,用量化阶 q 对直流系数 DC_{ij} 量化,整数型块均值 $m_{i,j}$ 和量化后整数型块均值 $m'_{i,j}$ 满足 $|m_{i,j}-m'_{i,j}|\leqslant\lceil\frac{q}{16}\rceil$,$\lceil x\rceil$表示大于等于实数 x 的最小整数;当 $q\leqslant32$ 时,$|m_{i,j}-m'_{i,j}|\leqslant2$。

证明 由上面定义及DCT变换的性质可知:

$$m_{i,j}=Round(\tilde{m}_{i,j}) \tag{2}$$

$$\tilde{m}_{i,j}=m_{i,j}+\Delta'_{i,j};\ -1/2\leqslant\Delta'_{i,j}<1/2 \tag{3}$$

$$m_{i,j}=\tilde{m}_{i,j}-\Delta'_{i,j} \tag{4}$$

$$DC_{i,j}=8\tilde{m}_{i,j} \tag{5}$$

其中,$Round(\cdot)$ 表示四舍五入取整函数,$\Delta'_{i,j}$ 表示取整误差。由量化的定义可得:

$$DC_{i,j}=Round\left(\frac{DC_{i,j}}{q}\right)\times q+\Delta''_{i,j};\quad -q/2\leqslant\Delta''_{i,j}<q/2 \tag{6}$$

由式(5)、(6)可知:

$$\tilde{m}_{i,j}=\frac{1}{8}Round\left(\frac{DC_{i,j}}{q}\right)\times q+\frac{1}{8}\Delta''_{i,j}=Round\left(\frac{1}{8}Round\left(\frac{DC_{i,j}}{q}\right)\times q\right)+\Delta'''_{i,j}+\frac{1}{8}\Delta''_{i,j}=m'_{i,j}+\Delta'''_{i,j}+\frac{1}{8}\Delta''_{i,j} \tag{7}$$

其中:$m'_{i,j}=Round\left(\frac{1}{8}Round\left(\frac{DC_{i,j}}{q}\right)\times q\right)$为量化后整数型块均值,$\Delta''_{i,j}$ 表示量化误差,$\Delta'''_{i,j}$ 表示取整误差,$-1/2\leqslant\Delta'''_{i,j}<1/2$。由式(4)和(7)容易得到:

$$|m_{i,j}-m'_{i,j}|=\left|\frac{1}{8}\Delta''_{i,j}+\Delta'''_{i,j}-\Delta'_{i,j}\right|\leqslant\left|\frac{1}{8}\Delta''_{i,j}\right|+|\Delta'''_{i,j}-\Delta'_{i,j}|\leqslant\frac{q}{16}+|\Delta'''_{i,j}-\Delta'_{i,j}| \tag{8}$$

由于 $|\Delta'''_{i,j}-\Delta'_{i,j}|<1$,$|m_{i,j}-m'_{i,j}|$为正整数,所以可得:

$$|m_{i,j}-m'_{i,j}|\leqslant\lceil\frac{q}{16}\rceil \tag{9}$$

其中$\lceil x\rceil$表示大于等于实数 x 的最小整数。当 $q\leqslant32$ 时,式(9)可进一步写为:

$$|m_{i,j}-m'_{i,j}|\leqslant2 \tag{10}$$

证毕

推论1 在对图像进行JPEG压缩情况下,当质量因子 $QF\geqslant25$ 时,$|m_{i,j}-m'_{i,j}|\leqslant2$。

这是因为在JPEG压缩中,质量因子 $QF=50$ 时,直流系数的量化阶 $q=16$。当 $QF<50$ 时,其直流系数的量化阶 $q_1=16\times(50/QF)$,当 $QF\geqslant50$ 时,$q_1=16\times(2-0.02\times QF)$,由此不难得到当 $QF\geqslant25$ 时,$q_1\leqslant32$。

推论 2　在对图像进行 JPEG 压缩情况下,当质量因子 $QF \geqslant 50$ 时,$| m_{i,j} - m'_{i,j} | \leqslant 1$。

推论 3　在对图像进行 JPEG 压缩情况下,当质量因子 $QF \geqslant 17$ 时,$| m_{i,j} - m'_{i,j} | \leqslant 3$。

由这些推论可分别得到不同的图像预处理方法,本文的预处理方法主要基于推论 1。

2.2　图像特征量提取及预处理方法

设整数型块均值 $m_{i,j}$ 的二进制形式为:

$$m_{i,j} = a^7_{i,j}a^6_{i,j}a^5_{i,j}a^4_{i,j}a^3_{i,j}a^2_{i,j}a^1_{i,j}a^0_{i,j} \tag{11}$$

取其高 5 位作为图像块 $B_{i,j}$ 的特征量,记为:

$$\hat{B}_{i,j} = a^7_{i,j}a^6_{i,j}a^5_{i,j}a^4_{i,j}a^3_{i,j} \tag{12}$$

由推论 1 不难看出,只要对式(11)中的低两位作适当的修改,使 $m_{i,j}$ 在增加 2 或者减少 2 时不会向第 4 位 $a^3_{i,j}$ 进位或者借位,就可以保证 $\hat{B}_{i,j}$ 在质量因子 $QF \geqslant 25$ 的 JPEG 压缩情况下保持不变。具体图像预处理方法如下。

1)将 $m_{i,j}$ 的二进制形式的低 2 位 $a^1_{i,j}a^0_{i,j}$ 依据第 3 位 $a^2_{i,j}$ 的情况作如下修改,修改后结果记为 $\hat{m}_{i,j}$:000 改为 010,001 改为 010,110 改为 101,111 改为 101,其余的不变。

2)与对应块内每个像素减去 $\Delta m_{i,j} = m_{i,j} - \hat{m}_{i,j}$。

容易证明经过如上预处理后的图像的整数型块均值正好为 $\hat{m}_{i,j}$,所以经过上述预处理方法处理过的图像,在质量因子 $QF \geqslant 25$ 的 JPEG 压缩情况下,其特征量 $\hat{B}_{i,j}$ 不变。

2.3　预处理对图像质量产生影响的分析

由上面的图像预处理方法不难看出,当假设一幅图像的块均值服从均匀分布时,本文方法只对一半的块均值修改,对块内每个像素的改变量不会超过 2,且容易算出预处理对图像块均值的改变量的平均值为:

$$E(\Delta m_{i,j}) = E(m_{i,j} - \hat{m}_{i,j}) = \frac{1}{8} \times 2 + \frac{1}{8}(-2) + \frac{1}{8} \times 1 + \frac{1}{8}(-1) = 0 \tag{13}$$

由此我们看到这种预处理对绝大部分图像不会造成明显的失真,表 1 的仿真结果也说明了这一点。

文献[7]利用概率分析的方法给出一种图像预处理方法,即对块均值的后两位做修改:00 改为 01,11 改为 10,这样预处理后可抵抗 $QF \geqslant 50$ 的 JPEG 压缩。表 1 为两种方法比较的仿真结果,由此可看出本文预处理方法在对图像质量影响小于文献[7]的情况下,将抗 JPEG 压缩能力由文献[7]的 $QF \geqslant 50$ 提升到 $QF \geqslant 25$。表 1 是预处理方法对部分 256 × 256 图像产生影响的仿真结果,仿真中未见可觉察的块效应现象。

这里需要注意的是 JPEG 压缩及预处理可能会使某些图像像素值产生溢出,从而对认证产生影响,为了消除这种影响,可预先对原始图像 I 作仿射变换,使其值域处于区间 $[d, 255-d]$,由文献[7]知,一般情况 d 取 10 就能满足要求,这种处理不会对原图像引入可察觉的失真。

表 1　本文预处理方法与文献[7]方法对图像影响的比较

测试图像	本文方法预处理后图像的 *PSNR*/dB	文献[7] 方法预处理后图像的 *PSNR*/dB
Lena 图像	45.46	43.47
Baboon 图像	44.79	42.95
Barb 图像	45.43	43.69
Couple 图像	45.85	44.19

3　图像认证方案的原理及算法实现

本文以 256 × 256 灰度图像为例,其结果可方便地推广到一般的 8 w × 8 h 灰度图像。其原理框图如图 1、2 所示。

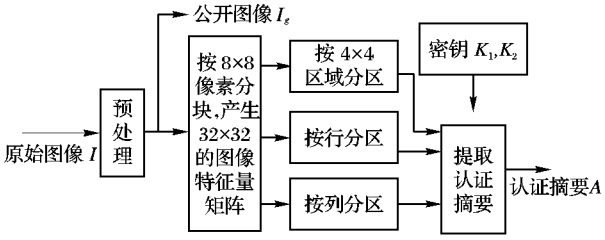


图 1　认证摘要的产生过程

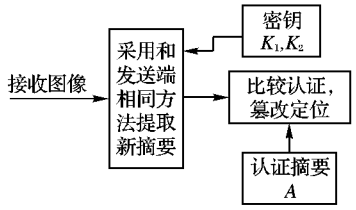


图 2　图像内容认证和篡改定位

3.1　认证摘要提取算法

第 1 步　将 256 × 256 的灰度图像按第 2 章的方法进行预处理。

第 2 步　将预处理后的灰度图像等分为 8 × 8 的块,取每个块的整数型块均值的高 5 位,形成 32 × 32 的图像特征量矩阵,表示为 $\boldsymbol{A}_{32\times32} = [a_{i,j}]$,其中 a_{ij} 为每个 8 × 8 块的整数型块均值的高 5 位。

第 3 步　对矩阵 $\boldsymbol{A}_{32\times32}$ 按三种方式进行区域划分:

1)等分为 4 × 4 的区域,取区域内每个 a_{ij},可得 4 × 4 × 5 = 80 比特,将其通过添 0 的方法扩展为 127 比特(也可以添加 47 比特随机二进制数,以增加算法的安全性,但要作为密钥),对每个区域都依次操作,可得:

$$B_i = (b_{1,i}, b_{2,i}, b_{3,i}, \cdots, b_{127,i}); \quad i = 1,2,\cdots,64 \tag{14}$$

2)按列等分为 32 × 1 的区域,取区域内每个 a_{ij},可得 32 × 1 × 5 = 160 比特,将其通过添 0 的方法扩展为 255 比特(也可以添加 95 比特随机二进制数,以增加算法的安全性,但要作为密钥),对每个区域都依次操作,可得:

$$R_j = (r_{1,j}, r_{2,j}, r_{3,j}, \cdots, r_{255,j}); \quad j = 1,2,\cdots,32 \tag{15}$$

3)按行等分为 1 × 32 的区域,同 2)可得:

$$L_k = (l_{1,k}, l_{2,k}, l_{3,k}, \cdots, l_{255,k}); \quad k = 1,2,\cdots,32 \tag{16}$$

第 4 步　给定 7 × 127 及 8 × 255 的二进制矩阵 $\boldsymbol{H}_1$ 和 $\boldsymbol{H}_2$,随机行向量 $\boldsymbol{M}$ 和 $\boldsymbol{N}$:

$$\boldsymbol{H}_1 = [1,2,3,\cdots,127] \tag{17}$$

$$\boldsymbol{H}_2 = [1,2,3,\cdots,255] \tag{18}$$

$$\boldsymbol{M} = (m_1, m_2, m_3, m_4, m_5, m_6, m_7) \tag{19}$$

$$\boldsymbol{N} = (n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8) \tag{20}$$

将 $\boldsymbol{H}_1$ 中的每一位写成 7 位二进制表示的列向量形式,即为 7 × 127 二进制矩阵;将 $\boldsymbol{H}_2$ 中的每一位写成 8 位二进制表示的列向量形式,即为 8 × 255 的二进制矩阵;一般常取其各种不同列置换矩阵。将 $\boldsymbol{H}_1$, $\boldsymbol{M}$ 作为密钥 K_1; $\boldsymbol{H}_2$, $\boldsymbol{N}$ 作为密钥 K_2。

第 5 步　采用第 1.1 节的方法,利用给定的 $\boldsymbol{H}_1$、$\boldsymbol{M}$ 对 B_i 提取摘要;利用给定的 $\boldsymbol{H}_2$、$\boldsymbol{N}$ 对 R_j 及 L_k 提取摘要。即:

$$\underline{\alpha_i} = \Gamma(B_i, \boldsymbol{M}, \boldsymbol{H}_1);$$

$$\underline{\alpha_j} \in \{0,1,2,\cdots,127\}; i = 1,2,\cdots,64 \tag{21}$$

$\underline{\beta}_j = \Gamma(R_j, N, H_2)$;

$\underline{\beta}_j \in \{0,1,2,\cdots,255\}; j = 1,2,\cdots,32$ (22)

$\underline{\delta}_k = \Gamma(L_k, N, H_2)$;

$\underline{\delta}_k \in \{0,1,2,\cdots,255\}; k = 1,2,\cdots,32$ (23)

其中$\underline{\alpha}_j$、$\underline{\beta}_j$和$\underline{\delta}_k$为认证摘要，通过秘密通道传送。容易计算认证摘要的二进制表示的总长度为$7\times64+8\times32\times2=960$比特。

当只对单区域篡改进行认证时，不需要提取4×4的区域的认证摘要，这时的认证摘要的二进制表示的总长度为$8\times32\times2=512$比特。

3.2 图像认证算法

第1步 摘要重新提取。对接收到的图像利用密钥K_1和K_2重复摘要提取算法的第2～5步，得到相应的$\underline{\alpha}_j'$、$\underline{\beta}_j'$和$\underline{\delta}_k'$。

第2步 篡改粗检测定位。通过秘密通道接收认证摘要$\underline{\alpha}_j$、$\underline{\beta}_j$和$\underline{\delta}_k$并与$\underline{\alpha}_i'$、$\underline{\beta}_j'$和$\underline{\delta}_k'$进行比较：

1）当$\underline{\alpha}_j=\underline{\alpha}_i'$时，第$i$个区域没有篡改，否则第$i$个区域被篡改；

2）当$\underline{\beta}_j=\underline{\beta}_j'$时，第$j$列没有篡改，否则第$j$列被篡改；

3）当$\underline{\delta}_k=\underline{\delta}_k'$时，第$k$行没有篡改，否则第$k$行被篡改。

第3步 篡改精确检测定位：

1）由第2步的1）实现区域篡改定位，假设定出篡改区域为B_i；

2）由第2步的2）、3）实现交叉定位，即如果$\underline{\delta}_k\neq\underline{\delta}_k'$且$\underline{\beta}_j\neq\underline{\beta}_j'$，则判定第$k$行和第$j$列的交叉区域被篡改，假设定出篡改区域为$R_j\cap L_k$；

3）取由1）、2）确定的篡改区域的交集作为真正的篡改位置即$B_i\cap R_j\cap L_k$。

这里的图像认证（篡改检测及定位）过程包括行认证、列认证及4×4区域（32像素×32像素）认证三个过程，取其认证结果的交集作为最终认证结果，将这种认证方法称为区域检测定位和交叉检测定位复合的篡改检测定位方法，该方法可以有效地解决多区域篡改图像的认证，具体分析详见第4.2节。

4 性能分析及实验结果

4.1 算法对JPEG压缩的鲁棒性

由第2章的理论分析可知，本算法提取的图像特征量对质量因子$QF\geqslant25$的JPEG压缩稳健，而认证摘要由图像特征量提取，所以本算法提取的行、列及区域数据的摘要对质量因子$QF\geqslant25$的JPEG压缩稳健。同时仿真实验也证明了这一点，实验验证结果如图3、4所示。图3(c)、4(c)为对预处理后图像先做$QF=25$的JPEG压缩，再做篡改的结果；图3(f)、4(f)为相应的篡改检测及定位结果，由此可以看出压缩没有对篡改检测和定位造成影响，即算法对JPEG压缩具有鲁棒性。综合第2章的结论，说明本文算法在对图像质量影响小于文献[7]的情况下，比文献[7]算法抗JPEG压缩能力强，且绝对稳健；同时在抗压缩方面也优于文献[6]，文献[6]只能抵抗$QF=80$的压缩。

4.2 多区域篡改情况下的篡改检测与定位能力分析

4.2.1 算法篡改检测概率分析

由于图像被篡改的区域都可以看作由若干个互不重叠的8×8的像素块组成，所以只要能检测到每个8×8的像素块是否被篡改，就可以实现篡改区域的检测和定位。本文的篡改检测概率就是指任意8×8的像素块的篡改检测概率。

对于任意8×8的像素块，本文算法是利用本文第1章提出的新方法从该像素块所在的行、列及区域的特征数据流中分别提取摘要，当特征数据流长度为n比特时，提取得摘要长度为$m=\mathrm{lb}(n+1)$比特，且由第1章的分析可知，当n比特的数据流发生改变时（由相关8×8的像素块的篡改所引起的改变），m比特摘要改变的概率为$P_y=(2^m-1)/2^m$；同时由于本文通过行、列及区域复合的方法进行篡改检测和定位，而这三者之间不一定独立。由此容易得到本文算法对256×256的灰度图像任意一个8×8的像素块的篡改检测概率为：

$$p_h\geqslant p_hp_lp_q=\frac{2^8-1}{2^8}\times\frac{2^8-1}{2^8}\times\frac{2^7-1}{2^7}\approx98\% \quad (24)$$

其中：p_h为行检测概率（$m=8$），p_l为列检测概率（$m=8$），p_q为区域检测概率（$m=7$）。

文献[7]的算法对于任意的8×8的像素块，先计算该像素块所在的行、列的特征数据流的Hash函数值，长度为128比特，再取其奇偶效验位来提取1比特的摘要，且当特征数据流发生改变时，这1比特摘要发生改变的概率为$P_h=1/2$。该算法采用行列交叉的方法进行篡改检测和定位，由于Hash函数的输出是伪随机的，也就是说，Hash函数生成的信息摘要是不可预见的，消息摘要看起来和原始的数据没有任何的关系，且原始数据的任何微小变化都会对生成的信息摘要产生很大的影响[1]，所以当8×8的像素块被篡改时，由此引起的行列摘要的变化可以看作相互独立的事件。由此不难得到文献[7]算法对256×256的灰度图像任意一个8×8的像素块的篡改检测概率为：

$$p_w=p_{wh}p_{wl}=0.5\times0.5=25\% \quad (25)$$

其中：p_{wh}为其行检测概率，p_{wl}为其列检测概率。

文献[6]的篡改检测概率也只有65%。

由以上分析不难看出，本文算法的篡改检概率优于文献[6-7]。文献[7]的算法以牺牲篡改检测概率为代价换取较短的认证摘要长度，虽然对256×256的灰度图像，只有64比特认证摘要，但篡改检测概率只有25%，限制了其实用价值。本文算法在保证98%以上的篡改检测概率的前提下，对单区域篡改情况认证摘要长度只有512比特，对多区域篡改情况只有960比特，而文献[7]的算法仅适合于单区域篡改认证，相比之下本文算法更具有实用价值。

4.2.2 多区域篡改情况下的篡改定位能力分析

行列交叉定位就是当某行摘要和列摘要都发生改变时判定行列交叉区域被篡改，是目前常用的一种篡改定位方法。当图像只有一个区域被篡改时该方法较为有效，而当图像有多个区域被篡改时，例如2个不在同一行或同一列的区域被篡改，该方法会定位出4个区域被篡改，从而出现严重的虚警，即没有篡改的地方被检测定位为有篡改。文献[2,7]都采用了这种方法，所以它无法解决多个篡改区域的正确检测和定位。而其他基于水印的认证方案[3-6]都是将一个块的摘要嵌入另一个不同块，所以也不能解决多区域篡改的认证。

本文算法在行列交叉检测定位方法的基础上，增加了区域检测，形成一种新的行、列及区域复合交叉的检测定位方法（见3.2节图像认证算法），这种方法可以最大限度地排除由纯交叉定位所引起的误判区域，完全适合多区域篡改图像的认证，仿真结果也证明了这一点。图3(c)为篡改图像，篡改区域为三个16×16的小块，图3(f)为本文算法的定位结果，图3(e)为交叉定位的结果；图4(c)为篡改图像，篡改是改变图中一小汽车的位置，图4(f)为本文算法的定位结果，

图 4(e)为交叉定位的结果。仿真结果表明,在多区域篡改的情况下,本文算法相对于文献[7]的交叉定位方法具有很高的定位精度和很低的检测定位虚警。

本文算法的定位精度由图像特征量的提取方法及定位方法决定,算法以 8×8 的图像块为单位提取特征量,采用行列交叉和区域复合的定位方法,由于取的是三者的交集,所以其定位精度和文献[2,6-7]相同,均为 8×8 像素。

4.3 算法的安全性分析

由于该算法采用了 K_1、K_2 方式的密钥,K_1 共有 127!×128 种选择,K_2 共有 255!×256 种选择,所以,攻击者很难采用穷举法获得密钥,从而无法得到正确的摘要。同时由于本算法属于签名认证,可有效抵抗块置换、拼贴及伪造攻击。

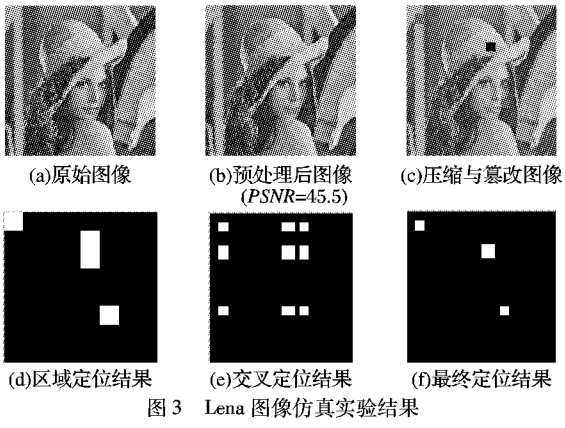
4.4 仿真实验结果

本文采用大量图像进行了多区域篡改认证实验,实验结果和性能分析所得的结论一致。由第 2 章的理论分析可知本文提取的特征量在 JPEG 压缩的情况下保持不变,所以对非压缩的篡改图像的认证结果和压缩后的篡改图像的认证结果是相同的,所以下面只给出压缩后的篡改图像的部分认证实验结果。

图 3 中篡改是在图像中随机插入大小为 16×16 像素的两个 CS 图片,一个大小为 16×16 像素的随机剪切;图 4 中篡改是改变图中一小汽车的位置;压缩图像指经质量因子 QF=25 的 JPEG 压缩后的图像。依据图像传输的一般原理,这里的压缩与篡改图像是指先压缩后篡改图像。图 3(b)、图 4(b)为预处理后图像,图像没有明显失真,且峰值信噪比较高(注意:这里的 $PSNR$ 是指仿射变换和预处理后图像的峰值信噪比);图 3(d)、图 4(d)为对图 3(c)、图 4(c)单纯的区域检测定位结果,结果表明,其定位精度较低,但检测概率高;图 3(e)、图 4(e)为对图 3(c)、图 4(c)单纯的交叉检测定位结果,结果表明,其定位精度较高,但虚警严重,会判定图像大面积被篡改;图 3(f)、图 4(f)为本文提出的行、列及区域复合交叉检测定位方法对图 3(c)、图 4(c)的检测定位结果,结果表明,检测定位准确,精度高,虚警很小;同时也表明,质量因子 $QF\geq 25$ 的 JPEG 压缩对篡改检测和定位未造成任何影响。

5 结语

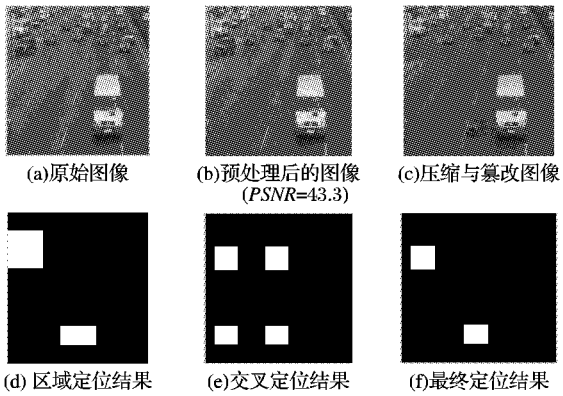
本文在研究摘要提取和图像预处理新方法的基础上,提出了一种基于数字签名的近似图像认证算法。该算法在抗 JPEG 压缩、篡改检测概率、虚警概率及篡改定位等方面都取得了良好的结果。在多区域篡改情况下,该算法具有一定的优势。该算法仅从图像块均值中提取摘要,对剪切、替换及位置改变等篡改认证效果良好,但是对细节性篡改认证有一定的局限性。在今后的工作中,可以将图像 DCT 变换的交流系数纳入认证摘要的提取范围,增强算法的认证能力。



(a)原始图像　(b)预处理后图像 ($PSNR$=45.5)　(c)压缩与篡改图像

(d)区域定位结果　(e)交叉定位结果　(f)最终定位结果

图 3　Lena 图像仿真实验结果



(a)原始图像　(b)预处理后的图像 ($PSNR$=43.3)　(c)压缩与篡改图像

(d) 区域定位结果　(e)交叉定位结果　(f)最终定位结果

图 4　交通汽车图像仿真实验结果

参考文献:

[1] 王育民, 刘建伟. 通信网的安全—理论与技术[M]. 西安: 西安电子科技大学出版社, 2002: 233-263.

[2] XIE LIE-HUA, ARCE G R, GRAVEMAN R F. Approximate image message authentication codes[J]. IEEE Transactions on Multimedia, 2001, 3(2): 242-252.

[3] 沃焱, 韩国强, 张波. 基于特征的静态图像内容认证方法[J]. 中国图象图形学报, 2006, 11(7): 1036-1042.

[4] 董刚, 张良, 张春田. 一种半脆弱性数字图像水印算法[J]. 通信学报, 2003, 24(1): 33-38.

[5] 钟桦, 焦李成. DCT 域半易损水印技术[J]. 计算机学报, 2005, 28(9): 1549-1557.

[6] 余淼, 和红杰, 张家树. 一种高定位精度的安全 JPEG 图像认证水印算法[J]. 中国科学: E 辑, 2007, 37(2): 315-328.

[7] 钟桦, 焦李成. 一种用于图像内容鉴别的数字签名方案[J]. 计算机学报, 2003, 26(6): 708-715.

[8] 王新梅, 肖国镇. 纠错码——原理与方法[M]. 西安: 西安电子科技大学出版社, 2003: 242-316.

(上接第 1616 页)

Java 智能卡安全策略中的应用,解决复杂情况下,如多个应用间非线性共享接口调用关系,如何验证安全策略是否被满足,安全策略如何更新,以及处理新安装 Applet 对 Java 智能卡内部环境造成的改变等问题。

参考文献:

[1] Sun Microsystems. Virtual machine specification Java card platform, Version 2.2.2 [EB/OL]. [2008-10-21]. http://java.sun.com/products/javacard/specs.html.

[2] Sun Microsystems. Runtime environment specification Java card platform, Version 2.2.2 [EB/OL]. [2008-10-21]. http://java.sun.com/products/javacard/specs.html.

[3] Sun Microsystems. Application programming interface Java card platform, Version 2.2.2 [EB/OL]. [2008-10-21]. http://java.sun.com/products/javacard/specs.html.

[4] GIRAND P. Which security policy for multiapplication smart card? [EB/OL]. [2008-10-10]. http://www.gemplus.com/smart/rd/publications/ps/Gir99sec.ps.

[5] BIEBER P, CAZIN J, GIRAND P, *et al*. Checking secure interactions of smart card applets: extended version [J]. Journal of Computer Security, 2002, (10)4: 369-398.

[6] DONG C Y, RUSSELLO G, DULAY N. Trust transfer in distributed systems [EB/OL]. [2008-10-10]. http://www.doc.ic.ac.uk/~cd04/papers/itrust07.pdf.