基于分块压缩感知的图像半脆弱零水印算法

赵春晖 [1] 刘 巍 [1]

摘 要 针对数字图像的内容认证和完整性保护问题, 提出了一种基于分块压缩感知 (Compressive sensing, CS) 的图像半脆弱零水印算法 (Block compressive sensing based image semi-fragile zero-watermarking, BCS-SFZ). 首先将图像划分成若干分块, 分块大小可以根据水印数据量和篡改定位精度调整. 再按照压缩感知理论对各个图像块进行观测, 并将观测值作为零水印信息注册保存. 实验结果表明, BCS-SFZ 算法可以准确定位非法篡改并借助水印信息恢复被篡改的区域. 压缩感知理论的引入为算法提供了保密性支持, 并且有利于实现图像成像与水印生成的同步, 同时该算法实现简单, 计算复杂度低.

关键词 半脆弱水印, 零水印, 压缩感知, 篡改定位, 图像恢复

引用格式 赵春晖, 刘巍. 基于分块压缩感知的图像半脆弱零水印算法. 自动化学报, 2012, **38**(4): 609−617

DOI 10.3724/SP.J.1004.2012.00609

Block Compressive Sensing Based Image Semi-fragile Zero-watermarking Algorithm

ZHAO Chun-Hui[1] LIU Wei[1]

Abstract To aim at content authentication and integrity protection problems in digital image application, a block compressive sensing based image semi-fragile zero-watermarking (BCS-SFZ) algorithm is proposed. Firstly, the image is divided into several sub blocks. The block size can be adjusted according to the data quantity of watermark and tamper localization accuracy. Secondly, each image block is observed in accordance with the theory of compressive sensing. Finally, the measurements are registered and preserved as the zero-watermarking. The experimental results show that BCS-SFZ can accurately locate the illegal tampering and recover the tampered area with the watermark information. The introduction of compressive sensing theory provides secrecy guarantee for the algorithm, and is conducive to synchronization of imaging and generating watermark. Also, the algorithm is simple and has a low computational complexity.

Key words Semi-fragile watermarking, zero-watermarking, compressive sensing (CS), tamper localization, image reconstruction

Citation Zhao Chun-Hui, Liu Wei. Block compressive sensing based image semi-fragile zero-watermarking algorithm. *Acta Automatica Sinica*, 2012, **38**(4): 609−617

数字水印作为一种有效解决图像版权、内容认证和完整性保护等问题的技术, 长期受到国内外学者和工程技术人员的关注. 但是传统的水印技术需要向宿主图像中嵌入水印, 影响了图像的原有信息, 这在很多对图像内容要求很高的应用中是不允许的[1]. 感知哈希 (Perceptual Hash)[2] 能够将数据集单向映射到感知摘要集中, 利用该技术可以在不引起图像感知偏差的情况下实现对图像内容真实性和完整性的认证. 但是哈希函数提取数字摘要的过程是不可逆的, 所以其提供的数字摘要只能用于篡改检测而不能恢复被篡改内容. "零水印" 技术能够可逆地提取图像特征构造水印, 不修改原始宿主图像, 有效地解决了不可感知性和水印嵌入之间的矛盾[3], 同时其提取的特征信息又可以用来指导篡改区域的恢复. 实现零水印的关键是: 1) 如何构造既包含有效特征又比原图像数据量少的水印数据; 2) 如何保证从相似图像中生成的水印不会互相混淆; 3) 如何保证水印的保密性, 即一旦水印信息泄露不会导致图像重要信息被攻击者获取利用.

近年来提出的压缩感知 (Compressive sensing, CS) 理论[4−6] 为零水印的构造提供了一种可行的解决途径. 首先, CS 观测值作为原图像的完整描述, 涵盖了其所有的特征并且数据量远少于原图像, 符合零水印实现的第一个条件; 而图像间的细微差别都会被观测值放大这一特征又保证了相近图像不会生成相似的观测值, 解决了水印混淆的问题; 另外,

收稿日期 2011-03-04 录用日期 2011-08-30
Manuscript received March 4, 2011; accepted August 30, 2011
国家自然科学基金 (61077079), 高等学校博士学科点专项基金 (20102304110013), 哈尔滨市优秀学科带头人基金 (2009RFXXG034) 资助
Supported by National Natural Science Foundation of China (61077079), Specialized Research Fund for the Doctoral Program of Higher Education (20102304110013), and Program Excellent Academic Leaders of Harbin (2009RFXXG034)
本文责任编委 黄继武
Recommended by Associate Editor HUANG Ji-Wu
1. 哈尔滨工程大学信息与通信工程学院 哈尔滨 150001
1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001

CS 观测值所具有的计算保密性 (Computational secrecy) 保证了攻击者在不具备密钥先验知识的情况下无法通过观测值推测原始图像的内容[7], 满足了零水印在保密性方面的要求. 与现有其他零水印算法相比, 基于 CS 的零水印算法对原始图像的信息提取更加全面、安全性更好. 值得一提的是, CS 理论的引入还有助于实现图像成像与水印生成的同步完成, 这种资源上的节省对于高带宽传感器等特定应用中减少计算消耗具有十分重要的意义.

针对数字图像的内容认证和版权保护问题, 本文提出了一种基于分块压缩感知的图像半脆弱零水印算法 (Block compressive sensing based image semi-fragile zero-watermarking, BCS-SFZ). 首先将图像划分成若干分块, 分块大小可以根据水印数据量和定位精度调整. 再对各个图像块进行分块 CS, 并将观测值作为零水印信息注册保存. 实验结果表明, BCS-SFZ 算法可以准确定位非法篡改, 并借助水印信息恢复被篡改的区域. CS 理论的引入为算法提供了保密性支持, 并且有利于实现图像成像与水印生成的同步完成, 同时该算法实现简单, 计算复杂度低.

1 压缩感知理论

1.1 压缩感知

考虑一个实值一维离散信号 $\boldsymbol{f} \in \mathbf{R}^N$, 可以用一组正交基 Ψ 表示为

$$\boldsymbol{f} = \Psi \boldsymbol{x} \tag{1}$$

其中, $\boldsymbol{x} \in \mathbf{R}^N$ 是正交变换的系数向量. 如果 $\sup(\boldsymbol{x}) = \{i : x_i \neq 0\}$, 那么当 $\sup(\boldsymbol{x}) \leq K$ 时称信号 $\boldsymbol{f}$ 为 K-稀疏信号. 在不进行变换的情况下直接通过一个与 Ψ 不相关的 $M \times N\,(M \ll N)$ 矩阵 Φ 获取信号 $\boldsymbol{f}$ 的观测值

$$\boldsymbol{y} = \Phi \boldsymbol{f} \tag{2}$$

根据式 (1) 可将式 (2) 等价为

$$\boldsymbol{y} = \Phi \boldsymbol{f} = \Phi \Psi \boldsymbol{x} = \Theta \boldsymbol{x} \tag{3}$$

称 $M \times N$ 矩阵 Φ 为观测矩阵, $N \times N$ 矩阵 Ψ 为稀疏分解矩阵. 压缩感知矩阵 $\Theta = \Phi\Psi$ 是一个矩阵. 可以证明[8] 解最小化范数优化方程

$$\hat{\boldsymbol{x}} = \min \| \boldsymbol{x}' \|_1, \qquad \Theta \boldsymbol{x}' = \boldsymbol{y} \tag{4}$$

可以准确恢复 K-稀疏信号. 当 Θ 满足受限等距特性 (Restricted isometry property, RIP)[9−10] 时可以保证长度为 N 的 $\boldsymbol{x}$ 可以由 M 个观测值, 通过式 (4) 精确恢复出来[11], 其中

$$M \geq \mathrm{O}(K \log N) \tag{5}$$

1.2 分块压缩感知

希望取得一幅图像的 M 个观测值, 在分块 CS 中整幅图像被分成大小为 $B \times B$ 的小块, 用 $\boldsymbol{x}_i$ 表示第 i 块的向量形式, 则相应的观测值 $\boldsymbol{y}_i$ 可以表示成

$$\boldsymbol{y}_i = \Phi_B \boldsymbol{x}_i \tag{6}$$

其中, Φ_B 是一个 $m \times B^2$ 矩阵, $m = \lfloor MB^2/N \rfloor$. 对整幅图像来说, 观测矩阵 Φ 是一个具有如式 (7) 形式的块对角矩阵. 值得注意的是, 在分块 CS 算法中仅需存储 $m \times B^2$ 矩阵 Φ_B, 而不需存储 $M \times N$ 矩阵 Φ[12], 显然这样可以极大地节省计算时需要的存储空间, 更加有利于算法的快速实现.

$$\Phi = \begin{bmatrix} \Phi_B & & & \\ & \Phi_B & & \\ & & \ddots & \\ & & & \Phi_B \end{bmatrix} \tag{7}$$

2 BCS-SFZ 算法

图像作为对成像场景的一个采样, 其像素数量的下限是受采样定律限制的. 而 CS 突破了采样定律的限制, 可以进一步通过线性投影的方式将图像映射成一个更为简洁的形式. 由于图像中包含的信息在映射中没有损失, 所以映射得到的观测值在信息量上与原图像是等价的, 并且是唯一对应的. 换句话说, 观测值就是原始图像所有特征的综合表示. 正是由于 CS 在特征提取中表现出的全面性, 促使我们考虑将其与零水印技术相结合.

2.1 水印的构造

本文提出的 BCS-SFZ 算法首先将原始图像分块, 再将各图像子块经过 CS 处理得到观测值, 最后将观测值组合在一起生成水印. 水印的构造流程如图 1 所示.

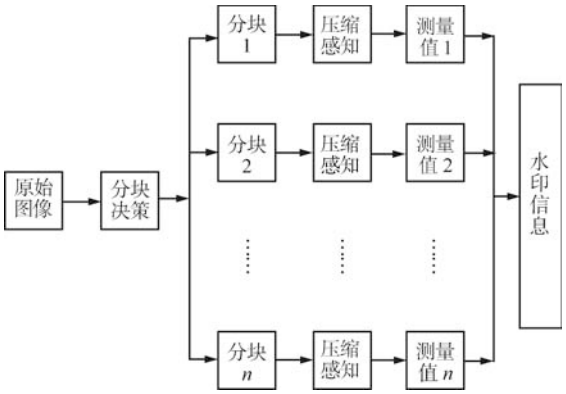


图 1 水印构造流程图

Fig. 1 Watermark forming process

图 1 中分块决策部分根据期望生成的水印尺寸和对篡改的定位精度制定分块策略. 考虑一幅大小为 $I_r \times I_c$ 的图像 I, 总像素数 $N = I_r I_c$. 将图像分成大小为 $B \times B$ 的小块 (假设 I_r 与 I_c 均能被 B 整除), 则图像总共被分成 $S = N/B^2$ 块. 为了后续讨论方便, 按照图 2 所示规则对各图像块进行统一编号. 在实际应用中, 分块尺寸越大则水印数据量就越少, 对篡改的定位精度越低, 对篡改部分的恢复计算量越大; 当分块尺寸减少时, 篡改定位的精度随之增加, 恢复篡改部分的计算量减少, 生成的水印数据量增加.

1	2	…	k
$k+1$	$k+2$	…	$2k$
⋮	⋮	⋱	⋮
$S-k+1$	$S-k+2$	…	S

图 2　图像块的编号规则

Fig. 2　Image block numbering rules

若要构造数据量为 M 的水印, 可以通过 CS 对每个图像子块各取 $m = \lfloor MB^2/N \rfloor$ 个观测值. 用 $\boldsymbol{Y}_i$ $(i = 1, 2, \cdots, S)$ 表示第 i 个图像块的观测值,

$$\boldsymbol{Y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{im} \end{bmatrix} \tag{8}$$

再将所有图像子块的观测值组合到一起生成最终的水印 $\boldsymbol{Y}_{wm}$, $\boldsymbol{Y}_{wm}$ 可以表示为

$$\boldsymbol{Y}_{wm} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \\ \vdots \\ \boldsymbol{Y}_S \end{bmatrix} \tag{9}$$

由于采用了 CS 理论, 所以水印的生成仅是一个简单的线性投影过程. 水印生成的时间复杂度取决于观测矩阵的大小. 根据前面设置的参数, 每一个图像块的特征提取所需的时间复杂度都是 $\mathrm{O}(mB^2)$, 整幅图像具有 $S = N/B^2$ 个分块, 所以特征提取的总时间复杂度为 $\mathrm{O}(mN)$.

生成 CS 观测矩阵的种子和分块尺寸可以作为水印密钥保存. 将水印 $\boldsymbol{Y}_{wm}$ 和密钥注册到相关内容认证和完整性保护数据库中, 并获得一个唯一的注册 ID 号, 该 ID 号与图像信息、水印密钥信息以及所有人 (或组织) 信息均登记在案, 如此图像即处在水印技术的保护之下了. 在需要使用注册水印信息时, 可以根据注册 ID 号进行检索查询. 图像检测流程具体如下: 当图像所有人 (或组织) 需要仲裁检测时, 需将其持有的 ID 号和密钥信息与管理数据库的专门机构中登记的信息进行核对, 核对无误后利用水印和密钥对待检测图像进行 (整体或局部) 检测, 当检测出篡改时可以对该篡改进行认定并作为法律依据, 必要时利用注册信息对图像内容进行恢复.

2.2　水印的检测

水印的检测算法是水印构造算法的逆过程. 在需要认证图像内容的真实性时, 首先, 取得密钥信息, 确定分块策略并生成分块 CS 的观测矩阵. 然后, 按照图 1 中水印生成步骤操作取得待测图像的水印, 将其与相关注册数据库中的注册水印比较, 确定图像内容的真实性. 具体流程如图 3 所示.

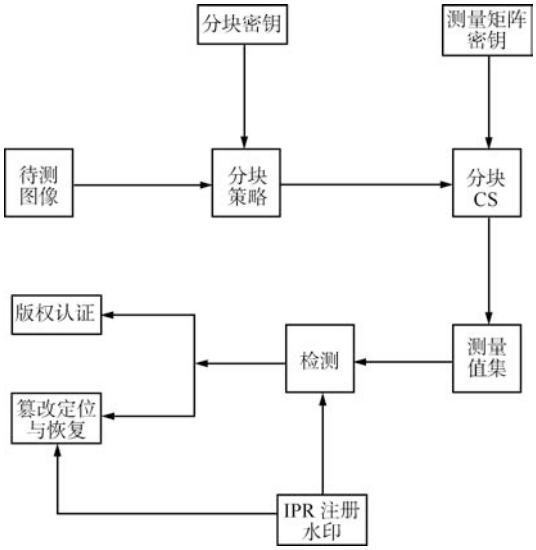


图 3　水印检测流程图

Fig. 3　Watermark detection flow chart

如果将待测图像中检测出的水印表示为

$$\boldsymbol{Y}_{wm}{}' = \begin{bmatrix} \boldsymbol{Y}_1{}' \\ \boldsymbol{Y}_2{}' \\ \vdots \\ \boldsymbol{Y}_S{}' \end{bmatrix} \tag{10}$$

其中, $\boldsymbol{Y}_i{}'$ $(i = 1, 2, \cdots, S)$ 表示待测图像中第 i 个

图像子块的观测值

$$\boldsymbol{Y}_i{}' = \begin{bmatrix} y_{i1}{}' \\ y_{i2}{}' \\ \vdots \\ y_{im}{}' \end{bmatrix} \tag{11}$$

为了计算水印之间的差异程度, 根据类似方差的思想, 可以将 $\boldsymbol{Y}_{wm}{}'$ 与原始图像在相关注册数据库中注册的水印 $\boldsymbol{Y}_{wm}$ 之间的欧氏距离的平方作为水印相似度的衡量标准

$$D_i = \sum_{k=1}^{n} (y_{ik}{}' - y_{ik})^2 \tag{12}$$

设定一个阈值 Th, 当 $D_i > Th$ 时认为待测图像中第 i 块被恶意篡改, 反之认为第 i 块内容真实.

本文在选择阈值时遵循下面准则. 考虑 Th 应该由两部分组成

$$Th = t_1 + \alpha t_2 \tag{13}$$

其中, t_1 表征合法操作强度, t_2 表征非法篡改可以引起的最小扰动, α 表示一个与 t_1 有关的强度系数. 本文中采用相似度量中的最小值作为 t_1, 选择一个常数 C_2 作为 t_2. 选择 α 时满足这样的原则: 当 t_1 为 0 时, $\alpha = 1$, α 随 t_1 增加而增加, 则有

$$t_1 = \min_{D_i}(D) \tag{14}$$

$$t_2 = C_2 \tag{15}$$

$$\alpha = \log\left(\frac{t_1 + 1}{C_1}\right) + 1 \tag{16}$$

其中, $C_1 = C_2 = 5 \times 10^4$ 为经验常数.

2.3 篡改恢复

在恢复被篡改的图像块时, 从注册水印中选择该图像块对应那部分观测值, 通过式 (4) 采用适当的最优化方法重建该图像块即可. 本文通过对像素梯度进行最优化, 修正式 (4) 为全变差最小化 (Total variation, TV) 方法. 令 $\| I \|_{\mathrm{TV}}$ 表示二维图像 I 的全变差, 若图像的每一个像素为 $I(t_1, t_2)$, $0 \le t_1, t_2 \le 0$, 则

$$\| I \|_{\mathrm{TV}} = \sum_{t_1, t_2} \sqrt{|D_1 I(t_1, t_2)|^2 + |D_2 I(t_1, t_2)|^2} \tag{17}$$

其中, D 为有限差分,

$$D_1 I = I(t_1, t_2) - I(t_1 - 1, t_2) \tag{18}$$

$$D_2 I = I(t_1, t_2) - I(t_1 - 1, t_2 - 1) \tag{19}$$

为了从观测值 y 中重建图像 I, 式 (4) 可改写为

$$I' = \min \| I \|_{\mathrm{TV}}, \qquad \Phi \boldsymbol{f} = \boldsymbol{y} \tag{20}$$

通过式 (20) 即可精确重建原始图像.

2.4 篡改恢复

CS 理论的意义在于其解放了采样定律对于硬件资源的限制, 只要一个信号在一组基下具有稀疏表示, 那么该信号就可以在不必经过稀疏变换的情况下, 仅通过与这组基不相关的另一组基上的少量线性投影重建. 采用这种方式获取信号极大地节省了传感器的数量, 可以借助 CS 的这种特性使用远少于图像精度要求的传感器进行成像. 文献 [13] 中介绍的压缩感知相机可以实现这一过程, 其成像原理如图 4 所示.

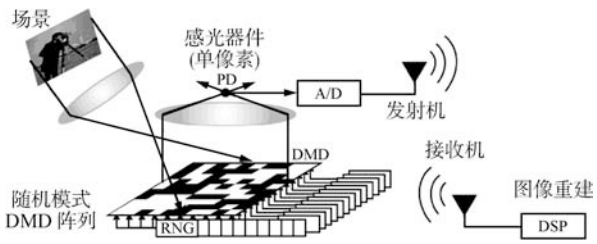


图 4 压缩感知相机原理图

Fig. 4 Compressive sensing camera block diagram

本文提出的 BCS-SFZ 算法利用分块 CS 的观测值作为半脆弱水印, 所以 CS 的观测过程与水印的构造过程是一致的. 将图 4 压缩感知相机中数字微镜系统 (Digital micromirror device, DMD) 的结构稍加修改就可以实现图像成像与水印生成的同步完成. 对压缩感知相机的修改本质上就是将原本的整体成像修改成分块成像, 修改后的 DMD 的工作机制如图 5 所示.

依然假设成像分辨率 $N = I_r I_c$, 图像分块大小为 $B \times B$ (假设 I_r 与 I_c 均能被 B 整除), 分块总数为 $S = N/B^2$. 按照图 2 所示的编号规则, 若当前要对图像的第 $k+2$ 块成像, 则随机数生成器 (Random number generator, RNG) 只控制 DMD 中与第 $k+2$ 块对应位置的小镜面翻转, 其他小镜面一律转向不朝向 (指令 0) 感光器件的方向. 对每个图像块进行 $m = \lfloor MB^2/N \rfloor$ 次观测, 即可获得 M 个总观测值, 同时也就构造出了数据量为 M 的水印.

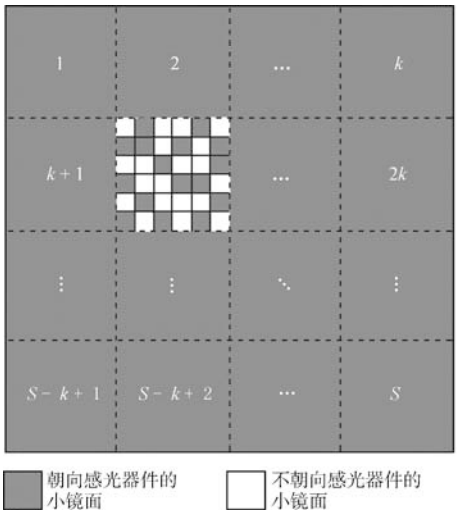


图 5　修改后的 DMD 原理图

Fig. 5　Modified DMD schematic diagram

3　水印检测原理及保密性分析

3.1　水印检测原理

采用 CS 观测值构造的零水印之所以能够检测篡改是因为其具有 "差异放大" 能力, 即图像间的细微差别都会被观测值放大. 参考式 (2), 将 $\boldsymbol{f}$ 投影到矩阵 Φ 上可以得到观测值 $\boldsymbol{y}$, 那么当 $\boldsymbol{f}$ 受到轻微扰动 $\Delta\boldsymbol{f}$ 影响时, 原信号变为 $\boldsymbol{f}' = \boldsymbol{f} + \Delta\boldsymbol{f}$, 此时其观测值也会受到相应的影响变为如下形式:

$$\boldsymbol{y}' = \Phi\boldsymbol{f}' = \Phi\boldsymbol{f} + \Phi\Delta\boldsymbol{f} = \boldsymbol{y} + \Delta\boldsymbol{y} \tag{21}$$

根据式 (2) 和式 (21) 可以看出

$$\Delta\boldsymbol{y} = \Phi\Delta\boldsymbol{f} \tag{22}$$

原信号受到的扰动对观测值所造成的影响就是 $\Delta\boldsymbol{f}$ 在观测矩阵上的投影 $\Delta\boldsymbol{y}$. CS 观测值所具有的 "差异放大" 能力实际就是通过观测矩阵元素对差异进行加权累加实现的. 为了分析方便, 先假设原信号 $\boldsymbol{f}$ 中只有一个位置发生了扰动, 即 $\Delta\boldsymbol{f}$ 中只有一个非零元素

$$\Delta\boldsymbol{f} = [0, 0, \cdots, \Delta\boldsymbol{f}_x, \cdots, 0] \tag{23}$$

此时计算观测值扰动 $\Delta\boldsymbol{y}$

$$\Delta\boldsymbol{y} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1x} & \cdots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2x} & \cdots & \phi_{2N} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \cdots & \phi_{Mx} & \cdots & \phi_{MN} \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \Delta f_x \\ \vdots \\ 0 \end{bmatrix} = \Delta f_x \begin{bmatrix} \phi_{1x} \\ \phi_{2x} \\ \vdots \\ \phi_{Mx} \end{bmatrix} = \begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_M \end{bmatrix} \tag{24}$$

Δy_i 的第 i 个元素可以表示为

$$\Delta y_i = \Delta f_x \cdot \phi_{ix} \tag{25}$$

可见原信号局部的轻微改变通过观测矩阵的投影造成了观测值全局的改变. 根据式 (12) 计算受到扰动前后观测值改变的剧烈程度

$$D = \sum_{i=1}^{M}(y_i' - y_i)^2 = \sum_{i=1}^{M}\Delta y_i^2 = \sum_{i=1}^{M}(\Delta f_x \cdot \phi_{ix})^2 = \Delta f_x^2 \sum_{i=1}^{M}\phi_{ix}^2 \tag{26}$$

在 M 取值较大的情况下 (可以看成 $M \to \infty$), 观测矩阵每列元素的分布与观测矩阵的整体分布相同, 可以根据观测矩阵的性质计算出 $\sum_{i=1}^{M}\phi_{ix}^2$ 的大小. 以满足 $\mathrm{N}(0,1)$ 的高斯矩阵和 0-1 矩阵为例:

1) 当使用高斯观测矩阵时, Φ 中元素满足 $\phi_{ij} \sim \mathrm{N}(0,1)$, 数学期望 $\mathrm{E}(\phi) = 0$, 方差 $D(\phi) = 1$. 于是 $D(\phi) = \mathrm{E}[(\phi - \mathrm{E}(\phi))^2] = \mathrm{E}(\phi^2) = 1$, 所以当 $M \to \infty$ 时, $\sum_{i=1}^{M}\phi_{ix}^2 = M \cdot \mathrm{E}(\phi^2) = M$, 可以得到 $D = \Delta f_x^2 \sum_{i=1}^{M} \cdot \phi_{ix}^2 = M\Delta f_x^2$.

2) 当使用 0-1 观测矩阵时, Φ 中元素取值 $P(\phi_{ij} = 0) = P(\phi_{ij} = 1) = 0.5$, 可知 $\mathrm{E}(\phi^2) = 0.5$. 当 $M \to \infty$ 时, $\sum_{i=1}^{M}\phi_{ix}^2 = M \cdot \mathrm{E}(\phi^2) = M/2$, 可以得到 $D = \Delta f_x^2 \sum_{i=1}^{M} \cdot \phi_{ix}^2 = M\Delta f_x^2/2$.

由此可以直观地看出, 在检测过程中原信号某一位置的扰动都会被以平方加权的方式放大. 以此类推不难得出如下结果, 当扰动 $\Delta\boldsymbol{f}$ 具有普适的形式 $\Delta\boldsymbol{f} = [\Delta f_1, \Delta f_2, \cdots, \Delta f_N]^{\mathrm{T}}$ 时, 根据式 (24) 观测值扰动 $\Delta\boldsymbol{y}$ 可表示为如下形式:

$$\Delta\boldsymbol{y} = \Delta f_1 \begin{bmatrix} \phi_{11} \\ \phi_{21} \\ \vdots \\ \phi_{M1} \end{bmatrix} + \Delta f_2 \begin{bmatrix} \phi_{12} \\ \phi_{22} \\ \vdots \\ \phi_{M2} \end{bmatrix} + \cdots + \Delta f_N \begin{bmatrix} \phi_{1N} \\ \phi_{2N} \\ \vdots \\ \phi_{MN} \end{bmatrix} \tag{27}$$

此时 $\Delta\boldsymbol{y}$ 的第 i 个元素可以表示为

$$\Delta y_i = \sum_{j=1}^{N} \Delta f_j \cdot \phi_{ij} \tag{28}$$

原信号受到的扰动以加权累加和的形式对观测值向量中所有元素造成影响. 根据式 (12) 计算受到扰动前后观测值改变的剧烈程度

$$\begin{aligned} D = & \sum_{i=1}^{M}(y_i' - y_i)^2 = \sum_{i=1}^{M} \Delta y_i^2 = \\ & \sum_{i=1}^{M}(\sum_{j=1}^{N} \Delta f_j \cdot \phi_{ij})^2 = \\ & \sum_{j=1}^{N}\sum_{i=1}^{M} \Delta f_j^2 \cdot \phi_{ij}^2 \end{aligned} \tag{29}$$

扰动对观测值的影响经过取平方后再一次被累加增强, 由此实现了观测值对信号差异的放大. 由于这种放大是以指数级数体现的, 所以当 Δf_j 较小时, D 也相对较小, 但随着 Δf_j 增大, D 的增加程度明显加快. 当图像受到 JPEG 压缩等合法操作时, 图像的视觉效果不会发生明显改变, 此时 Δf_j 较小接近 0, 所以 D 较小; 而当图像受到篡改时, 图像的视觉内容发生了改变, 所以 Δf_j 往往很大, 若干个较大的 Δf_j 产生的影响经过叠加后使 D 形成了峰值. 所以可通过设定一个合理的阈值有效地区分合法操作与非法篡改.

3.2 水印保密性分析

压缩感知过程本身可以看成是一个对信号加密的过程, 而且这种加密可以在不必负担独立加密协议带来的额外计算消耗的情况下, 同时提供信号压缩和加密保证. 加密过程中观测矩阵可以被看成是完全可靠的一次性密钥. 文献 [7] 中通过定理 1 及推论 1 证明了不具备密钥先验知识的攻击者只能通过穷举的方式逐个尝试密钥的所有可能形式直到他认为自己重建出了原始图像为止, 这是一个 NP-hard 问题, 如果攻击者不知道密钥的生成范围的话, 那么他需要尝试的可能性就有无穷多个; 当攻击者尝试错误的密钥时他将以概率 1 重建出一个无任何实际视觉意义的错误图像.

攻击者对于零水印图像的攻击分为三种情况: 1) 仅知道图像内容; 2) 仅知道水印内容; 3) 既知道图像内容又知道水印内容. 几乎所有的水印系统都是根据第 1) 种情况设计的, 攻击者很难在不知道水印构造的前提下对水印图像进行不被察觉的篡改. 第 2) 种情况针对含有保密内容的图像, 例如军事应用中的遥感图像以及用于法庭作证的证据图像, 攻击者在无法窃取原始图像的情况下从水印数据库中窃取了零水印信息, 并希望由此推断出原始图像的内容. 水印的保密性由压缩感知观测值的保密性提供, 观测矩阵中元素的随机性决定了观测值具有随机噪声的性质, 独立于密钥使用时不具备任何实际意义, 这一特性使得攻击者在不经过重建的前提下无法通过分析观测值来推断原始图像的内容, 确保了第 2) 种情况下的保密性. 对于第 3) 种情况, 参考第 3.1 节中水印检测原理可知, 任何对于图像具有实际意义的修改都会造成像素灰度值较大范围的变化, 而这种变化均会被水印检测出来, 这就使得攻击者在同时获得图像与水印的情况下也无法构造出既不会被检测又具有实际意义的篡改. 当然如果攻击者具备同时替换版权人手中的密钥以及数据库中的注册水印的能力的话, 那么他可以对图像进行有效篡改, 而在此种情况下所有水印算法对其均失效, 本文不做进一步的讨论.

4 实验结果及效果分析

实验中选取如图 6 所示的尺寸为 512 像素 $\times$ 512 像素的灰度图像进行测试. BCS-SFZ 算法中采用的分块大小为 $B = 32$, CS 观测矩阵统一采用 256 像素 $\times$ 1 024 像素的高斯随机矩阵. 对图 6 所示的原始图像中的内容进行篡改, 篡改后得到的待测图像如图 7 所示. 详细的篡改信息列于表 1.

表 1 篡改详细信息

Table 1 Details of tampering operations

被篡改块	篡改像素数	最大改变量	平均改变量
第 99 块	30	132	52
第 100 块	275	149	66
第 121 块	223	160	43
第 122 块	421	162	44
第 123 块	117	168	35

图 6 原始图像

Fig. 6 Original image

图 7　篡改图像

Fig. 7　Tampered image

采用 BCS-SFZ 算法对待测图像进行水印检测, 检测结果如图 8 所示, 图 9 将待测图像中检测到的被篡改图像子块与原始图像对应子块进行了放大对比. 对照图 8 和图 9 可以清楚地看到水印可以将篡改有效地放大, 使之与未受篡改的图像块区分开.

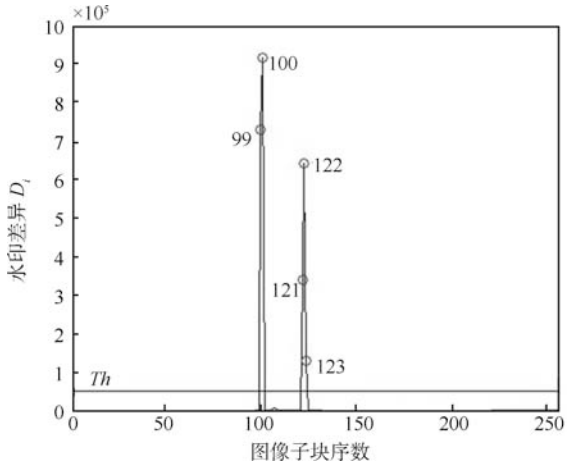
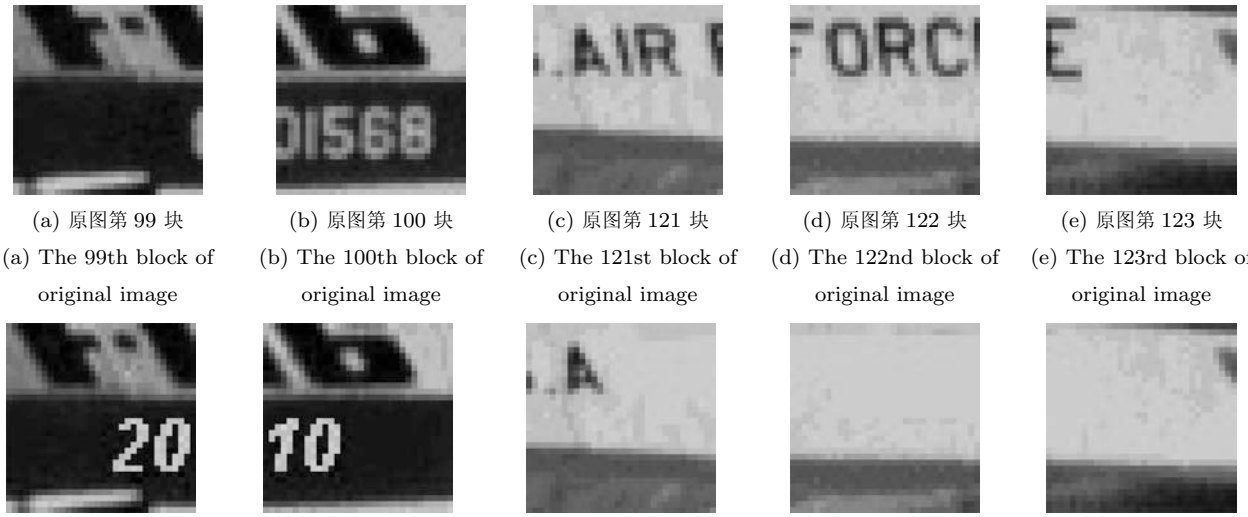


图 8　水印检测结果

Fig. 8　Watermark detection results



(a) 原图第 99 块
(a) The 99th block of original image

(b) 原图第 100 块
(b) The 100th block of original image

(c) 原图第 121 块
(c) The 121st block of original image

(d) 原图第 122 块
(d) The 122nd block of original image

(e) 原图第 123 块
(e) The 123rd block of original image

(f) 待测第 99 块
(f) The 99th block of tested image

(g) 待测第 100 块
(g) The 100th block of tested image

(h) 待测第 121 块
(h) The 121st block of tested image

(i) 待测第 122 块
(i) The 122nd block of tested image

(j) 待测第 123 块
(j) The 123rd block of tested image

图 9　篡改位置对照

Fig. 9　Tampered position control

表 2　合法操作与非法篡改检测

Table 2　Legal operations and illegal tamper detection

操作类型	最大检测值 $D_{\max}$	最小检测值 $D_{\min}$	阈值 Th	篡改检测结果
JPEG $Q=80$	6.6546E+004	5.8013E+004	1.4652E+005	无篡改
JPEG $Q=80$ + 篡改	9.2144E+005	5.6405E+004	1.4417E+005	第 99, 100, 121, 122, 123 块被篡改
JPEG $Q=60$	7.7690E+004	5.9406E+004	1.4856E+005	无篡改
JPEG $Q=60$ + 篡改	9.3323E+005	5.8462E+004	1.4718E+005	第 99, 100, 121, 122, 123 块被篡改
JPEG $Q=40$	9.7429E+004	6.2264E+004	1.5271E+005	无篡改
JPEG $Q=40$ + 篡改	9.4601E+005	5.8587E+004	1.4736E+005	第 99, 100, 121, 122, 123 块被篡改
JPEG $Q=20$	1.3112E+005	6.2320E+004	1.5279E+005	无篡改
JPEG Q= 20 + 篡改	9.5767E+005	6.5619E+004	1.5753E+005	第 99, 100, 121, 122, 123 块被篡改

BCS-SFZ 算法对图像的合法操作具有较强的鲁棒性, 能够抵抗诸如 JPEG 压缩等操作, 并且可以将其与恶意篡改有效区分开来. 将图 6 所示原始图像经过不同品质因数 JPEG 压缩处理之后进行检测, 再对压缩后的图像进行如图 7 所示的篡改之后进行检测, 检测结果列于表 2 中. 在品质因数 $Q = 80$ 和 $Q = 20$ 的情况下经过篡改的图像检测效果和未经过篡改的图像检测效果分别如图 10～图 13 所示. 从图中可以看到 JPEG 压缩对于水印的影响表现为水印差异 D_i 不同程度的增加, 但其程度远没有恶意篡改的影响明显, 可以简单地通过设定阈值 Th 加以区分.

对比图 10～图 13 可知, BCS-SFZ 算法对于局部恶意篡改具有非常明显的放大效果, 而对于 JPEG 压缩造成的差异却不十分敏感, 这种特性保证了水印半脆弱性的实现.

由于 BCS-SFZ 算法利用了第三方的优势, 恢复水印不会由于数据的压缩而去除或者被破坏; 再加上 CS 理论的有力支撑, 使得被篡改部分可以被精确恢复. 在检测到篡改的情况下, 根据被篡改位置的编号取得对应位置的原始水印, 按照 TV 最小化方法重建该图像块, 并替换被篡改部分. 对图 6 所示原始图像进行品质因数 $Q = 20$ 的 JPEG 压缩, 再对压缩后的图像进行如图 7 所示的篡改, 此时其峰值信噪比 (Peak signal to noise ratio, PSNR) 为 25.1062, 效果如图 14 所示. 对图 14 进行恢复, 得到如图 15 所示的恢复效果 PSNR = 25.0369, 从图中可以看出被篡改信息得到了有效地恢复.

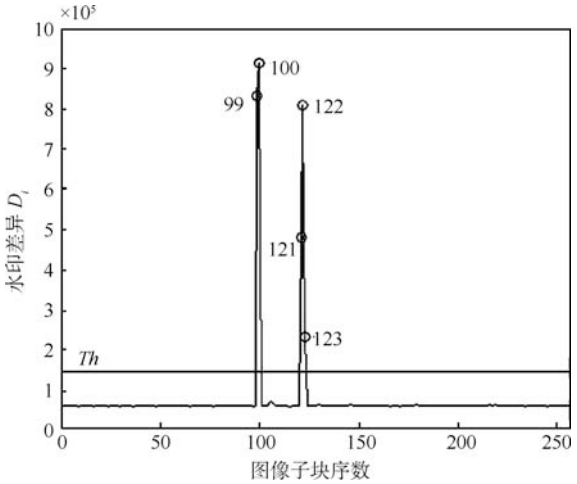


图 10 有篡改情况下 $Q = 80$ 的检测结果

Fig. 10 Detection results of $Q = 80$ with tampering

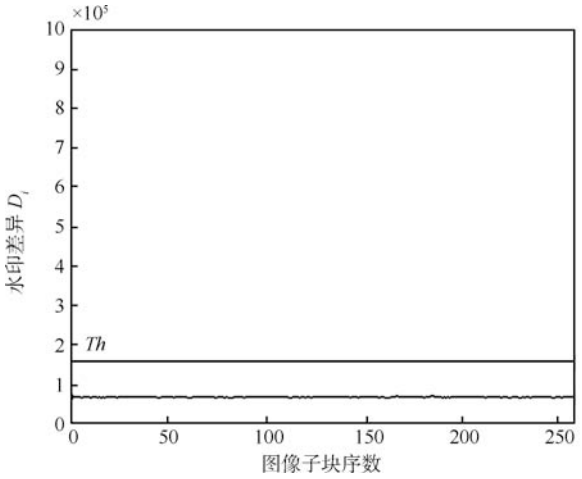


图 11 无篡改情况下 $Q = 80$ 的检测结果

Fig. 11 Detection results of $Q = 80$ without tampering

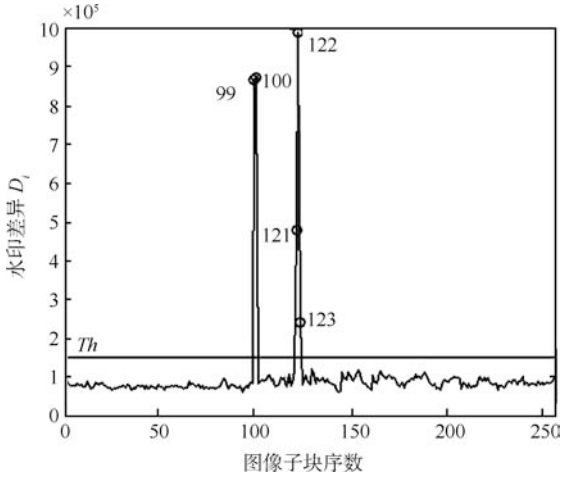


图 12 有篡改情况下 $Q = 20$ 的检测结果

Fig. 12 Detection results of $Q = 20$ with tampering

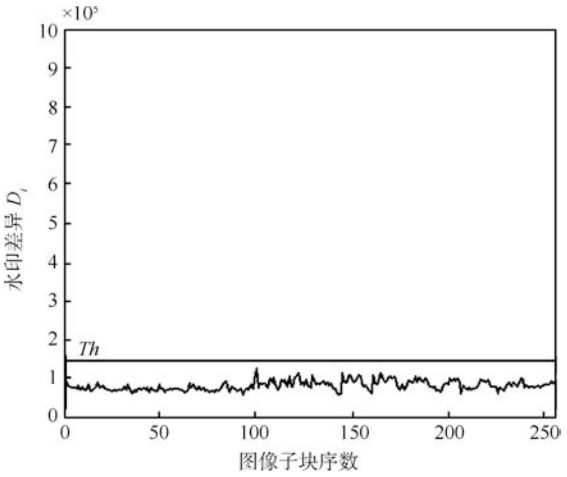


图 13 无篡改情况下 $Q = 20$ 的检测结果

Fig. 13 Detection results of $Q = 20$ without tampering

图 14　恢复前效果

Fig. 14　Effect before rehab

图 15　恢复后效果

Fig. 15　Effect after rehab

5　结论

本文提出了一种基于分块压缩感知的图像半脆弱零水印算法 (BCS-SFZ), BCS-SFZ 将压缩感知理论与零水印技术相结合, 实现了一种新的半脆弱水印策略. 图像通过压缩感知获得的观测值包含原图像携带的所有信息, 可以简洁而全面地表征图像的全部特征. 以观测值作为水印具有操作简单 (仅需要通过一个线性变换即可)、对恶意篡改敏感和对合法操作容忍度高、保密性好等优点, 而且被篡改内容可以通过水印信息恢复. BCS-SFZ 采用了零水印技术来保证原始图像数据不被干扰, 满足了对图像质量要求较高的应用.

References

1 Zhang Li-Bao, Ma Xin-Yue, Chen Qi. Image zero-watermarking algorithm based on region of interest. *Journal on Communications*, 2009, **30**(S2): 117−120
(张立保, 马新悦, 陈琪. 基于感兴趣区的图像零水印算法. 通信学报, 2009, **30**(S2): 117−120)

2 Niu Xia-Mu, Jiao Yu-Hua. An overview of perceptual Hashing. *Acta Electronica Sinica*, 2008, **36**(7): 1405−1411
(牛夏牧, 焦玉华. 感知哈希综述. 电子学报, 2008, **36**(7): 1405−1411)

3 Wen Quan, Sun Tan-Feng, Wang Shu-Xun. Concept and application of zero-watermark. *Acta Electronica Sinica*, 2003, **31**(2): 214−216
(温泉, 孙锬锋, 王树勋. 零水印的概念与应用. 电子学报, 2003, **31**(2): 214−216)

4 Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, **52**(4): 1289−1306

5 Candes E J, Wakin M B. An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition. *IEEE Signal Processing Magazine*, 2008, **25**(2): 21−30

6 Romberg J. Imaging via compressive sampling. *IEEE Signal Processing Magazine*, 2008, **25**(2): 14−20

7 Rachlin Y, Baron D. The secrecy of compressed sensing measurements. In: Proceedings of the 46th Annual Allerton Conference on Communication, Control and Computing. Illinois, USA: IEEE, 2008. 813−817

8 Donoho D L. For most large underdetermined systems of equations, the minimal L_1-norm near-solution approximates the sparsest near-solution. *Communications on Pure and Applied Mathematics*, 2006, **59**(7): 907−934

9 Baraniuk R, Davenport M, DeVore R, Wakin M. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 2008, **28**(3): 253−263

10 Candes E, Romberg J. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 2006, **6**(2): 227−254

11 Dossal C, Peyre G, Fadili J. A numerical exploration of compressed sampling recovery. *Linear Algebra and Its Applications*, 2010, **432**(7): 1663−1679

12 Gan L. Block compressed sensing of natural images. In: Proceedings of the15th International Conference on Digital Signal Processing. Cardiff, UK: IEEE, 2007. 403−406

13 Duarte M, Davenport M, Takhar D, Laska J N, Sun T, Kelly K F, Baraniuk R G. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 2008, **25**(2): 83−91

赵春晖　哈尔滨工程大学信息与通信工程学院教授. 主要研究方向为智能信息与图像处理, 非线性信号处理, 通信信号处理.
E-mail: zhaochunhui@hrbeu.edu.cn
(**ZHAO Chun-Hui**　Professor at the College of Information and Communication Engineering, Harbin Engineering University. His research interest covers intelligent information and image processing, nonlinear signal processing, and communication signal processing.)

刘　巍　哈尔滨工程大学信息与通信工程学院博士研究生. 2006 年获得哈尔滨工程大学信息与通信工程学院学士学位. 主要研究方向为非线性信号与图像处理. 本文通信作者.
E-mail: liuwei16@hrbeu.eu.cn
(**LIU Wei**　Ph. D. candidate at the College of Information and Communication Engineering, Harbin Engineering University. He received his bachelor degree from the College of Information and Communication Engineering, Harbin Engineering University in 2006. His research interest covers nonlinear signal and image processing. Corresponding author of this paper.)