doi:10.11959/j.issn.1000-436x.2017177

动态更新失真代价的自适应 JPEG 隐写算法

汤光明[1]，孙艺[1]，徐潇雨[1]，王宇[2]
(1. 中国人民解放军信息工程大学，河南 郑州 450001；2. 吉林省四平军分区，吉林 四平 136000)

摘 要：以往的自适应 JPEG 隐写算法大多只在秘密信息嵌入之前计算图像失真，无法在嵌入过程中动态地调节失真代价。考虑到信息嵌入时的交互影响，提出一种动态更新失真代价的自适应 JPEG 隐写算法。首先，分析量化步长、待嵌 DCT(discrete cosine transform)系数绝对值以及扰动误差 3 种影响嵌入波动的因素。进而，提出失真代价更新策略 EUS(embedding update strategy)，使图像的失真代价能够动态地更新。最后，结合此策略设计并实现一种自适应 JPEG 隐写算法。实验表明该算法能够显著提高 JPEG 隐写的安全性。

关键词：信息安全；自适应隐写；失真函数；JPEG
中图分类号：TP391 **文献标识码：**A

Adaptive JPEG steganography based on distortion cost updating

TANG Guang-ming[1], SUN Yi[1], XU Xiao-yu[1], WANG Yu[2]
(1. PLA Information Engineering University, Zhengzhou 450001,China;
2. Military Subarea of Siping in Jilin Province, Siping 136000,China)

Abstract: Previous adaptive JPEG steganography algorithms mostly calculate image distortion before secret information embedding, so they can't dynamically adjust distortion costs. Considering the mutual impacts during embedding, an adaptive JPEG steganography algorithm based on distortion cost updating was proposed. Firstly, three factors that affect embedding fluctuations were analyzed, including quantization step, the absolute value of the quantized DCT coefficient to be modified and perturbation error. Then the embedding update strategy (EUS) was proposed, which enabled to dynamically update the distortion costs. After that, an adaptive JPEG steganography algorithm was implemented combining the strategy. The experimental result illustrates that the algorithm can significantly improve the secure performance of JPEG steganography algorithm.

Key words: information security, adaptive steganography, distortion function, JPEG

1 引言

数字隐写是信息隐藏的一个重要分支，通过将秘密信息隐藏在公开的数字载体中，如文本、图像、音频、视频，并保持嵌入的秘密信息不被人所感知，以达到在传输合法载体的同时进行隐蔽通信的目的。JPEG 图像是目前应用最广泛的图像格式，基于 JPEG 图像的隐写算法随之成为研究的焦点。

为了提高 JPEG 隐写的安全性，不少学者做了大量的研究。近年来，基于载体内容特性的自适应隐写[1]逐渐成为研究的热点。自适应隐写基于最小化加性失真函数[2]，将失真定义为对各个载体元素进行信息嵌入的代价之和。其关键问题是失真函数的定义和隐写编码的设计。PQ(perturbed quantization) [3]隐写算法利用原始未压缩图像作为边信息，将秘密信息嵌入到量化取整误差接近 0.5 的 DCT 系数中，使嵌入误差接近标准 JPEG 压缩中的取整误差。MME(modified matrix encoding) [4] 算法采用了

收稿日期：2016-09-29；**修回日期：**2016-11-16
通信作者：孙艺，mickyfaith@163.com
基金项目：国家自然科学基金资助项目（No.61272488）
Foundation Item: The National Natural Science Foundation of China (No.61272488)

修正矩阵编码，在秘密信息嵌入时选择各系数组中失真最小的修改模式进行嵌入，使总体失真最小化。nsF5(no-shrinkage F5)[5] 算法避免了收缩效应，并且采用湿纸码进行嵌入，显著提高了嵌入效率。

病灶格子编码（STC, syndrome trellis code）[6]的提出使在加性失真下的秘密信息嵌入实现了理论上的上界值。因此，目前自适应隐写的研究重点为失真函数的设计。Wang[7] 用 DCT 图像块熵度量块复杂度，并用 STC 编码嵌入，减少了失真并提高了嵌入效率 UED（uniform embedding distortion）[8]和它的改进算法 UERD（uniform embedding revisited distortion）[9]引入了均匀嵌入思想，将嵌入修改均匀地分散在量化 DCT 系数中，显著提高了隐写的安全性。Huang[10]将包括直流系数在内的全部 DCT 系数按照优先级分成 2 组，分别赋予不同的失真代价，较好地保持了原始图像的统计特性。J-UNIWARD[11]和基于边信息的 SI-UNIWARD 利用方向滤波来定义图像嵌入失真，使图像噪声、纹理区域像素修改代价小，图像平滑区域像素修改代价大，提高了算法的抗隐写分析性能。Tomas[12]提出了一种通用策略，可以在最小化加性失真框架下将任何隐写算法与边信息结合，从而提高该算法的安全性能。以上算法在隐写安全性方面都有较好的表现。但这些算法只在秘密信息嵌入之前计算图像失真，无法在嵌入过程中动态地调节失真代价，忽略了信息嵌入过程中的交互影响，对影响失真的因素考虑得还不够全面，失真函数的定义仍有待完善。Li 等[13]提出了一种动态更新像素失真代价的 CMD 隐写算法，该算法根据待嵌像素四邻域像素的修改方式动态调节失真代价，提高了算法的抗检测性。

由于 DCT 系数具有块内和块间相关性，在信息嵌入过程中，当对一个系数进行修改时，与之相关的其他系数的统计特性会随之受到影响。因此，不能简单地在秘密信息嵌入之前一次性定义图像失真，而应该考虑信息嵌入过程中 DCT 系数间的交互影响。本文在 Li 等研究的基础之上，结合 JPEG 图像自身的系数特性，提出一种动态更新失真代价的 JPEG 隐写算法。首先，对原始图像预处理。将原始图像 DCT 系数划分成一系列子图像块，基于子图像块生成嵌入子块，同时，将秘密信息划分为与嵌入子块数目相对应的部分。之后，选取一种现有的初始加性失真函数(如 UERD 和 J-UNIWARD)对原始图像进行失真代价初始化，利用 STC 编码将秘密信息的第一部分嵌入到原始图像的第一个嵌入子块中。在对其余子块进行嵌入时，每完成一个子块的嵌入就进行一次失真代价更新。实验表明，该算法能够有效地减小嵌入波动并保持 DCT 系数相关性，在对抗各种隐写检测算法时有较高安全性。

2 STC 编码

STC[6]编码，又称校验格编码，是目前为止编码性能最好的隐写编码方法，能够使加性失真下的秘密信息嵌入达到理论上的上界值。STC 编码是一种特殊的矩阵编码，其主要思想是利用奇偶校验矩阵 $\boldsymbol{H}$ 建立图像网格，将最小化失真代价的问题转化为求最短路径的问题，而该问题可以通过维特比译码算法快速得到。其中，嵌入失真 $\rho(x,y)$ 可根据具体情况自行定义。奇偶校验矩阵 $\boldsymbol{H}$ 是由一系列特殊的子矩阵 $\hat{\boldsymbol{H}}$ 级联拼接而成的， $\hat{\boldsymbol{H}}$ 由共享密钥 K 随机产生，大小为 $h\times w$，其中，h 决定了编码的速度和效率，w 的值由嵌入率 α 决定。$\boldsymbol{H}$ 的构成方式如图 1 所示。

$$\hat{\boldsymbol{H}}=\begin{pmatrix}1&0\\1&1\end{pmatrix}\rightarrow \boldsymbol{H}=\begin{pmatrix}1&0&&&&&&&&\\1&1&1&0&&&&&&\\&&1&1&&&&&&\\&&&&\ddots&&&&&\\&&&&&&1&0&&\\&&&&&&1&1&1&0\end{pmatrix}$$

图 1 奇偶检验矩阵的构成

在具体应用中，利用 STC 编码进行秘密信息嵌入和提取的表达式为

$$Emb(\boldsymbol{X},m)=\arg\min_{\boldsymbol{Y}\in C(m)} G(\boldsymbol{X},\boldsymbol{Y}) \tag{1}$$

$$Ext(\boldsymbol{Y})=\boldsymbol{H}\boldsymbol{Y} \tag{2}$$

其中，$\boldsymbol{X}$ 为原始图像，$\boldsymbol{Y}$ 为载密图像，m 为秘密信息，$C(m)=\left\{z\in\{0,1\}^n\,|\,\boldsymbol{H}z=m\right\}$ 为 m 的陪集，$G(\boldsymbol{X},\boldsymbol{Y})$ 为秘密信息嵌入后的图像总失真。编码嵌入的目的是使总嵌入失真尽可能小，即在 $C(m)$ 中寻找与原始图像 $\boldsymbol{X}$ 有最小失真的载密图像 $\boldsymbol{Y}$。接收方收到载密图像 $\boldsymbol{Y}$ 后，将其与奇偶校验矩阵 $\boldsymbol{H}$ 相乘即可得到秘密信息 m。

3 失真代价动态更新策略的设计

现有的自适应 JPEG 隐写算法大多只在秘密信

息嵌入之前计算图像失真，忽略了信息嵌入时的交互影响，无法在嵌入过程中动态地调节失真代价。但在实际的嵌入过程中，当对一个 DCT 系数进行修改时，与之相关的其他系数的统计特性会随之受到影响。本节从 JPEG 图像相关性出发，综合分析影响 JPEG 图像信息嵌入波动的因素，使 DCT 系数的修改方向与其邻域系数嵌入波动总和的方向相一致，提出一种失真代价动态更新策略，从而在秘密信息嵌入过程中动态地调整 DCT 系数的失真代价。

3.1　影响信息嵌入波动的因素

DCT 域的分块操作使 JPEG 图像既有块内相关性又有块间相关性。块内相关性主要表现为中心 DCT 系数与 4 邻域 DCT 系数间的关系，块间相关性主要表现为中心 DCT 系数与 4 邻域子块内频率相同的 4 个 DCT 系数之间的关联。对于 JPEG 图像而言，在 DCT 系数的块内与块间邻域系数中，不同位置的 DCT 系数对嵌入波动的贡献是有差别的，需要考虑系数位置对嵌入波动的影响。本节主要考虑 3 个影响信息嵌入波动的因素，分别为量化步长（QS, quantization step）、待嵌量化 DCT 系数绝对值（VQ, absolute value of the quantized DCT coefficient to be modified）以及扰动误差（PE, perturbation error）。

1) 量化步长

JPEG 压缩是有损过程，量化是控制 JPEG 图像压缩率的关键一步。量化式为

$$X_{i,j}=\left[\frac{U_{i,j}}{Q_{i,j}}\right] \tag{3}$$

其中，$U_{i,j}$ 和 $X_{i,j}$ 分别为量化前和量化后的 DCT 系数，$Q_{i,j}$ 为量化矩阵中的元素，$[x]$ 为取整操作。图 2 为质量因子为 80 的标准量化表。

$$\begin{bmatrix} 6 & 4 & 4 & 6 & 10 & 16 & 20 & 24 \\ 5 & 5 & 6 & 8 & 10 & 23 & 24 & 22 \\ 6 & 5 & 6 & 10 & 16 & 23 & 28 & 22 \\ 6 & 7 & 9 & 12 & 20 & 35 & 32 & 25 \\ 7 & 9 & 15 & 22 & 27 & 44 & 41 & 31 \\ 10 & 14 & 22 & 26 & 32 & 42 & 45 & 37 \\ 20 & 26 & 31 & 35 & 41 & 48 & 48 & 40 \\ 29 & 37 & 38 & 39 & 45 & 40 & 41 & 40 \end{bmatrix}$$

图 2　质量因子为 80 的标准量化表

从图 2 可以看出，不同位置的 DCT 系数对应的量化步长不同，低频 DCT 系数对应的量化步长低于高频 DCT 系数。由 DCT 变换的原理可知，DCT 变换后，图像能量主要集中于直流系数和低频交流系数。可以推知，量化步长越小的 DCT 系数包含图像能量越高。因此，在量化步长较小的 DCT 系数上嵌入秘密信息造成的图像波动较大。

2) 待嵌量化 DCT 系数绝对值

图像能量不仅和量化步长有关，还与待嵌量化 DCT 系数绝对值的大小有关。DCT 系数的绝对值越大，该系数所含图像能量越高，秘密信息嵌入对原始图像造成的影响越大。

图 3 为一幅 JPEG 图像的一个 8×8 子块的量化 DCT 系数。从图 3 可知，绝对值较大的 DCT 系数大多集中在系数子块的低频区域，即量化 DCT 系数值的大小同样反映了图像能量的大小。从而建立了量化步长与待嵌量化 DCT 系数绝对值之间的关系，即 DCT 系数的绝对值与量化步长成负相关。在相同嵌入幅度下，在绝对值较大的 DCT 系数上嵌入秘密信息造成的图像波动较大。

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图 3　JPEG 图像一个子块的 DCT 系数

3) 扰动误差

若原始未压缩图像在发送方是可得到的，设秘密信息比特的长度为 M，$U_{i,j}^m$ 和 $X_{i,j}^m(i,j=1,2,\cdots,8, m=1,2,\cdots,M)$ 分别为量化前后的 DCT 系数。则取整误差（RE, rounding error）可表示为

$$RE=\left|U_{i,j}^m-X_{i,j}^m\right| \tag{4}$$

设 $Y_{i,j}^m$ 为嵌入秘密信息后的 DCT 系数，则嵌入误差（EE, embedding error）可表示为

$$EE=\left|Y_{i,j}^m-X_{i,j}^m\right| \tag{5}$$

扰动误差可表示为嵌入误差与量化误差的差值

$$PE=EE-RE=\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right| \tag{6}$$

在式（6）中，$PE\geqslant0$，原因在于若秘密信息嵌入在量化 DCT 系数的 LSB 位，那么 $Y_{i,j}^m$ 的取值可能为 $U_{i,j}^m$、$U_{i,j}^m+1$、$U_{i,j}^m-1$，这 3 种情况下 EE 值均大

于 *RE*。扰动误差反映了秘密信息嵌入对原始图像统计特性的影响，扰动误差越大，秘密信息嵌入造成的图像波动越大。

3.2 EUS 策略的设计

结合 3.1 节对影响信息嵌入波动因素的分析，本节提出了失真代价动态更新的策略。首先提出信息嵌入波动度的概念，对不同位置的 DCT 系数赋予不同的权重；之后根据邻域系数波动度的总和判断中心系数的修改方向，从而达到保持 DCT 系数相关性的目的。

给出信息嵌入波动度的定义如下。

定义 1 设原始图像为 $\boldsymbol{X}$，隐秘图像为 $\boldsymbol{Y}$，令 $\boldsymbol{D}=\boldsymbol{Y}-\boldsymbol{X}$ 为修改方式矩阵，$d_{i,j}$ 为 $\boldsymbol{D}$ 中的元素，表示对 DCT 系数 $x_{i,j}$ 的修改方式，对于三元隐写，$d_{i,j}\in\{+1,0,-1\}$。则对 DCT 系数 $x_{i,j}$ 进行嵌入的波动度为

$$\sigma_{i,j}=\frac{d_{i,j}\left|V_{i,j}\right|^{\varphi_1}\left(\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right|\right)}{q_{i,j}^{\varphi_2}} \tag{7}$$

其中，$\dfrac{\left|V_{i,j}\right|^{\varphi_1}\left(\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right|\right)}{q_{i,j}^{\varphi_2}}$ 为权重因子，用于衡量波动的大小。$q_{i,j}$、$\left|V_{i,j}\right|$ 和 $\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right|$ 分别表示与 d_{ij} 相对应位置的量化步长，量化 DCT 系数的绝对值以及扰动误差的值，φ_1、φ_2 为参数，由实验确定，参数确定的方法如下。

1) 随机从图像库中选取 2 500 幅图像，并设定一个固定的嵌入率，如 0.3（即每个交流 DCT 系数平均嵌入 0.3 bit 秘密信息）。

2) 选取 CC-JRM 隐写检测特征与集成分类器对以下取值范围内的参数值分别进行测试。

$$\varphi_1=\{0,0.1,0.2,\cdots,1\}$$
$$\varphi_2=\{0,0.1,0.2,\cdots,1\}$$

其中，φ_1和φ_2取值均小于 1 的原因是 $q_{i,j}$、$\left|V_{i,j}\right|$ 相比于 $\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right|$ 数值偏大，为了平衡三者之间的权重需要使φ_1和φ_2取值偏小。

3) 选取测试结果最优的参数值。最终选定 $\varphi_1=0.4,\varphi_2=0.5$。

设 DCT 系数 $x_{i,j}$ 与其块内和块间邻域系数嵌入波动度的总和为 W，为保持秘密信息嵌入后 DCT 系数间的相关性，应使 W 的值达到最小，如式（8）（以块内邻域系数为例，块间同理）所示。

$$\begin{aligned}\min W=&\left|\sigma_{i-1,j}-\sigma_{i,j}\right|+\left|\sigma_{i,j-1}-\sigma_{i,j}\right|+\\&\left|\sigma_{i,j+1}-\sigma_{i,j}\right|+\left|\sigma_{i+1,j}-\sigma_{i,j}\right|\\\Rightarrow&(\sigma_{i-1,j}-\sigma_{i,j})^2+(\sigma_{i,j-1}-\sigma_{i,j})^2+\\&(\sigma_{i,j+1}-\sigma_{i,j})^2+(\sigma_{i+1,j}-\sigma_{i,j})^2\\=&4\sigma_{i,j}^2-2\sigma_{i,j}(\sigma_{i-1,j}+\sigma_{i,j-1}+\sigma_{i,j+1}+\sigma_{i+1,j})+\\&(\sigma_{i-1,j}^2+\sigma_{i,j-1}^2+\sigma_{i,j+1}^2+\sigma_{i+1,j}^2)\end{aligned} \tag{8}$$

式(8)中仅 $\sigma_{i,j}$ 为未知数，其他可视为常数。因此，当 W 取最小值时，$\sigma_{i,j}$ 的取值为

$$\sigma_{i,j}=\frac{1}{4}(\sigma_{i-1,j}+\sigma_{i,j-1}+\sigma_{i,j+1}+\sigma_{i+1,j}) \tag{9}$$

综合考虑块内与块间邻域系数得出 $\sigma_{i,j}$ 的最终取值为

$$\begin{aligned}\sigma_{i,j}=\frac{1}{4}(&\sigma_{i-1,j}+\sigma_{i,j-1}+\sigma_{i,j+1}+\sigma_{i+1,j}+\\&\sigma_{i-8,j}+\sigma_{i,j-8}+\sigma_{i,j+8}+\sigma_{i+8,j})\end{aligned} \tag{10}$$

由于权重因子 $\dfrac{\left|V_{i,j}\right|^{\varphi_1}\left(\left|Y_{i,j}^m-X_{i,j}^m\right|-\left|U_{i,j}^m-X_{i,j}^m\right|\right)}{q_{i,j}^{\varphi_2}}$ 为正值，对于三元隐写而言，为保持图像相关性，波动度 $\sigma_{i,j}$ 的符号应与修改方向一致，即当 $\sigma_{i,j}$ 的取值分别为正值、负值和 0 时，相应地，修改方式 $d_{i,j}$ 分别应选取为 +1、−1、0。

4 自适应隐写算法设计

结合第 3 节提出的失真代价更新策略，本节提出了一种自适应 JPEG 隐写算法，考虑了信息嵌入过程中的交互作用。算法的主要思想如下：首先，对原始图像进行预处理，将原始图像的 DCT 系数划分为一系列不重合的子图像块，基于子图像块生成嵌入子块，同时，将秘密信息划分为与嵌入子块数目相对应的部分；其次，利用初始加性失真函数（如 UERD 和 J-UNIWARD）将秘密信息的第一部分嵌入到原始图像的第一个嵌入子块中；之后，结合失真代价更新策略 EUS(embedding update strategy)，按照预定的扫描顺序分批次地更新剩余嵌入子块的失真代价并嵌入秘密信息。

4.1 秘密信息嵌入

秘密信息嵌入的具体过程如下。

Step1 将大小为 $n_1\times n_2$ 的原始图像 $\boldsymbol{X}$ 划分为大小为 $L_1\times L_2$ 的一系列不重合的子图像块。其中，$L_1,L_2\geqslant 1$。如图 4 所示，将一个 4×6 的矩阵划分成

互不重合的 2×3 的子图像块。

$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	$x_{1,6}$
$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	$x_{2,6}$
$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	$x_{3,5}$	$x_{3,6}$
$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	$x_{4,5}$	$x_{4,6}$

图 4　图像分块示意

Step2　基于子图像块生成图像嵌入子块 $\boldsymbol{S}_{a,b}$。如图 5 所示，在图 4 基础上生成 6 个嵌入子块 $\boldsymbol{S}_{1,1},\boldsymbol{S}_{1,2},\boldsymbol{S}_{1,3},\boldsymbol{S}_{2,1},\boldsymbol{S}_{2,2},\boldsymbol{S}_{2,3}$，计算式为

$$\boldsymbol{S}_{a,b}=\left\{(i,j)\middle|i=a+\tau_a L_1, j=b+\tau_b L_2\right\} \tag{11}$$

其中，$a=\{1,\cdots,L_1\}$，$\tau_a\in\left\{0,1,\cdots,\left\lfloor\frac{n_1}{L_1}\right\rfloor-1\right\}$，$b=\{1,\cdots,L_2\}$，$\tau_b\in\left\{0,1,\cdots,\left\lfloor\frac{n_2}{L_2}\right\rfloor-1\right\}$。

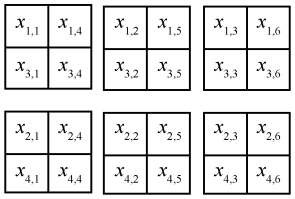


图 5　嵌入子块示意

Step3　将 M 个秘密信息比特分为 $L_1\times L_2$ 个部分，分配方式如下。

1) 根据不同的初始失真函数，在应用失真代价更新策略之前，先按照设定的嵌入率对原始图像进行预嵌入，并计算出每一嵌入子块中嵌入的秘密信息比特数量。

2) 将计算出的结果分配给各个嵌入子块，作为最终的嵌入比特数目。

Step4　确定嵌入子块的嵌入顺序。为方便起见，以水平 Z 字形为嵌入顺序，如图 6 所示。按照嵌入顺序将嵌入子块重新标记为 $\boldsymbol{S}_t$，$t\in\{1,2,\cdots,L_1\times L_2\}$。

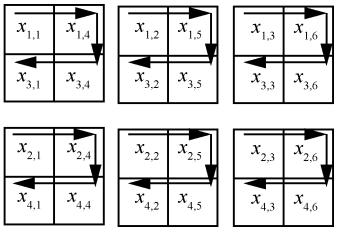


图 6　嵌入顺序示意

Step5　参数初始化。令 $t=1$，$\boldsymbol{Y}=\boldsymbol{X}$，修改方式矩阵 $\boldsymbol{D}=\boldsymbol{Y}-\boldsymbol{X}$，确定并计算如下初始加性失真函数。

1) 不基于边信息的 UERD[9]失真函数为

$$\rho_{ij}=\begin{cases}\dfrac{0.5(q_{(i+1)j}+q_{i(j+1)})}{D_{mn}+0.25\sum\limits_{d\in\hat{D}}d}, & (i,j)\bmod 8=(0,0)\\[2ex] \dfrac{q_{ij}}{D_{mn}+0.25\sum\limits_{d\in\hat{D}}d}, & \text{其他}\end{cases} \tag{12}$$

$$D_{mn}=\sum_{k=0}^{7}\sum_{l=0}^{7}|x_{kl}|q_{kl} \tag{13}$$

其中，x_{ij} 表示第 m 行、第 n 列的 DCT 系数块中的 DCT 系数，该块的块能量为 D_{mn}（避免直流系数的影响取 $x_{00}=0$）。q 为与 x 相对应的量化步长的值。$\hat{D}=\{D_{(m-1)(n-1)},D_{(m-1)n},D_{(m-1)(n+1)},D_{m(n-1)},D_{m(n+1)},D_{(m+1)(n-1)},D_{(m+1)n},D_{(m+1)(n+1)}\}$ 表示第 m 行、第 n 列的 DCT 系数块 8 邻域的块能量集合。

2) 基于边信息的 SI-UNIWARD[11]失真函数

$$\begin{aligned}D^{(\mathrm{SI})}(X,Y)&\triangleq D(P,Y)-D(P,X)\\&=\sum_{k=1}^{3}\sum_{u,v}\frac{\left|W_{uv}^{(k)}(P)-W_{uv}^{(k)}(Y)\right|-\left|W_{uv}^{(k)}(P)-W_{uv}^{(k)}(X)\right|}{\varepsilon+\left|W_{uv}^{(k)}(P)\right|}\end{aligned} \tag{14}$$

其中，$\boldsymbol{P}$ 为原始未压缩图像，$W_{uv}^{(k)}(\boldsymbol{X})$ 表示 $\boldsymbol{X}$ 在第 k 个分解方向滤波器下 (u,v) 位置上的小波系数值。ε 为控制参数，作用是防止分母为 0 导致计算式无意义。

Step6　设位于 (i,j) 的载体元素的失真代价为 $\rho_{i,j}=(\rho_{i,j}^{+},\rho_{i,j}^{-},\rho_{i,j}^{0})$，其中，$\rho_{i,j}^{+}$、$\rho_{i,j}^{-}$、$\rho_{i,j}^{0}$ 分别表示修改方式为 +1、−1、0 时的失真代价，易知 $\rho_{i,j}^{0}=0$。设初始失真函数为 F，$f_{i,j}$ 为 F 中的元素。按照以下方式计算或更新失真代价 $\rho_{i,j}$。

1) 若 $t=1$，令 $\rho_{i,j}^{+}=\rho_{i,j}^{-}=f_{i,j}$。

2) 否则，根据修改方式矩阵 $\boldsymbol{D}=\boldsymbol{Y}-\boldsymbol{X}$，更新失真代价。若初始失真函数不基于边信息，令

$$\rho_{i,j}^{+}=\begin{cases}\dfrac{f_{i,j}}{\alpha}, & \sum\limits_{d_{ia}\in N_{ia},d_{ir}\in N_{ir}}\left(\dfrac{d_{ia}\times|V_{ia}|^{\varphi_1}}{q_{ia}^{\varphi_2}}+\dfrac{d_{ir}\times|V_{ir}|^{\varphi_1}}{q_{ir}^{\varphi_2}}\right)>0\\ f_{i,j}, & \text{其他}\end{cases} \tag{15}$$

$$\rho_{i,j}^{+}=\begin{cases}\dfrac{f_{i,j}}{\alpha}, & \sum\limits_{d_{ia}\in N_{ia},d_{ir}\in N_{ir}}\left(\dfrac{d_{ia}|V_{ia}|^{\varphi_1}}{q_{ia}^{\varphi_2}}+\dfrac{d_{ir}|V_{ir}|^{\varphi_1}}{q_{ir}^{\varphi_2}}\right)<0\\ f_{i,j}, & \text{其他}\end{cases} \tag{16}$$

若初始失真函数基于边信息，令

$$\rho_{i,j}^{+}=\begin{cases}\dfrac{f_{i,j}}{\alpha}, & \sum\limits_{d_{ia}\in N_{ia},d_{ir}\in N_{ir}}(\sigma_{ia}+\sigma_{ir})>0\\ f_{i,j}, & \text{其他}\end{cases} \tag{17}$$

$$\rho_{i,j}^{-}=\begin{cases}\dfrac{f_{i,j}}{\alpha}, & \sum\limits_{d_{ia}\in N_{ia},d_{ir}\in N_{ir}}(\sigma_{ia}+\sigma_{ir})<0\\ f_{i,j}, & \text{其他}\end{cases} \tag{18}$$

其中，α 为控制参数。$N_{ia}=\{d_{i+1,j},d_{i-1,j},d_{i,j+1},d_{i,j-1}\}$ 表示 $d_{i,j}$ 的 4 邻域中 DCT 系数的修改方式。$N_{ir}=\{d_{i+8,j-8},d_{i-8,j-8},d_{i+8,j+8},d_{i-8,j+8}\}$ 表示 $d_{i,j}$ 的 4 邻域子块内相同频率位置 DCT 系数的修改方式。

在实际计算过程中，由于 d_{ir} 的取值均为 0，因此，用 d_{ir}' 的值近似代替 d_{ir}。对于任意 $d_{ir}\in N_{ir}$，令

$$d_{ir}'=\frac{1}{4}\sum_{d_{irr}\in N_{irr}}d_{irr} \tag{19}$$

其中，$N_{irr}=\{d_{i+1,r},d_{i-1,r},d_{i,r+1},d_{i,r-1}\}$ 表示 d_{ir} 的 4 邻域中 DCT 系数的修改方式。

Step7 根据更新后的失真代价，按照设定好的嵌入顺序分批次地在嵌入子块中用 STC 编码嵌入秘密信息。

Step8 重复此嵌入操作直至 $t=L_1\times L_2$。输出隐秘图像 $\boldsymbol{Y}$。

4.2 秘密信息提取

Step1 利用隐秘图像 $\boldsymbol{Y}$ 生成嵌入子块 $\boldsymbol{S}_t$，$t\in\{1,2,\cdots,L_1\times L_2\}$。

Step2 对于每一个嵌入子块，利用 $\boldsymbol{M}_t=\boldsymbol{H}\boldsymbol{S}_t^{\mathrm{T}}$ 提取出秘密信息，直至提取出所有秘密信息。

Step3 按照嵌入子块顺序将秘密信息组合起来，得到最终的原秘密信息 $\boldsymbol{M}$。

实际上，考虑到 JPEG 图像的 DCT 系数分块特性，L_1 和 L_2 的值均被设定为 8。

5 实验及安全性分析

本节首先对隐写策略的控制参数 α 进行确定，之后对算法的抗检测性能进行分析。

5.1 实验设置

实验基于 BOSSBASE[14]图像库，该图像库包含 10 000 幅尺寸为 512 像素×512 像素的图像。图像内容十分丰富，包括风景、人物、动植物等。将 10 000 幅图像进行 JPEG 压缩，质量因子分别为 75 和 95。对于每个质量因子，选取 5 000 幅用于训练，5 000 幅用于测试，在不同的嵌入率下（0.1~0.5）进行实验。秘密信息比特是随机产生的。本文提出的算法按以下方式命名：结合 EUS 策略与 UERD 初始失真函数的算法命名为 UERD-EUS，结合 EUS 策略与 SI-UNIWARD 初始失真函数的算法命名为 SI-UNIWARD-EUS。

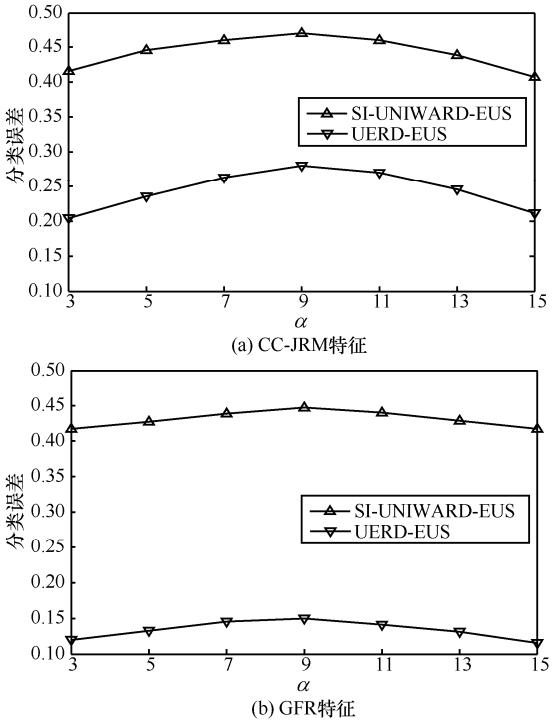
实验选取现有的一些有代表性的隐写算法与本文提出的算法进行比较，这些算法有 nsF5[5]、UERD[9]和 SI-UNIWARD[11]。此外，还选取了 CC-JRM[15]特征、DCTR[16]特征以及 GFR[17]特征进行隐写检测实验。由于集成分类器[18]能在高维特征下进行快速自动的训练，因此实验选取集成分类器进行分类测试。最小平均错误率 P_{E} 可表示为

$$P_{\mathrm{E}}=\min_{P_{\mathrm{FA}}}\frac{1}{2}(P_{\mathrm{FA}}+P_{\mathrm{MD}}P_{\mathrm{FA}}) \tag{20}$$

其中，P_{FA} 和 P_{MD} 分别表示虚警率和漏检率。P_{E} 越高表示算法的抗检测性能越好。

5.2 参数确定

为了得到控制参数 α 的最优值，实验在固定的质量因子（QF=75）及嵌入率（0.3）下，针对 UERD-EUS 与 SI-UNIWARD-EUS 算法，对一定范围内的 α 值用 CC-JRM 特征和 GFR 特征进行了测试，实验结果如图 7 所示。



(a) CC-JRM特征

(b) GFR特征

图 7 控制参数 α 在不同隐写检测特征下的测试结果

从图 7 可以看出，对于不同的初始失真函数和隐写检测特征，各种控制参数α下分类误差的峰值并不会有太大的变化。在α为3～15 范围内，当$\alpha=9$时，分类误差最大，因此，选取$\alpha=9$为最优控制参数。

5.3　抗检测性能分析

将 UERD-EUS 和 SI-UNIWARD-EUS 算法与 3 种现有的隐写算法进行比较实验，它们分别是非自适应隐写算法 nsF5[5]、自适应隐写算法 UERD[9]和 SI-UNIWARD[11]。在 75 和 95 这 2 种质量因子下，利用 CC-JRM 特征、DCTR 特征和 GFR 特征对以上算法进行对比实验，实验结果如图 8~图 10 所示。

从图中可知，除 CC-JRM 特征在 *QF*=95 时小嵌入率下的实验结果以外，SI-UNIWARD-EUS 的抗检测性能相对于其他隐写算法是最优的，SI-UNIWARD 的安全性表现稍弱于 SI-UNIWARD-EUS，并且这 2 种算法的安全性明显高于其他几种算法。可以看出，本文提出的 EUS 失真代价更新策略确实能够提高初始失真函数的安全性能。SI-UNIWARD 和 UERD 隐写算法抗检测性能好的原因可能在于这2种算法将秘密信息嵌入到所有 DCT 系数中，包括直流系数和零值交流系数，而传统隐写算法大多在非零交流 DCT 系数上嵌入信息，因此传统的隐写检测算法难以捕捉对嵌入操作敏感的特征。

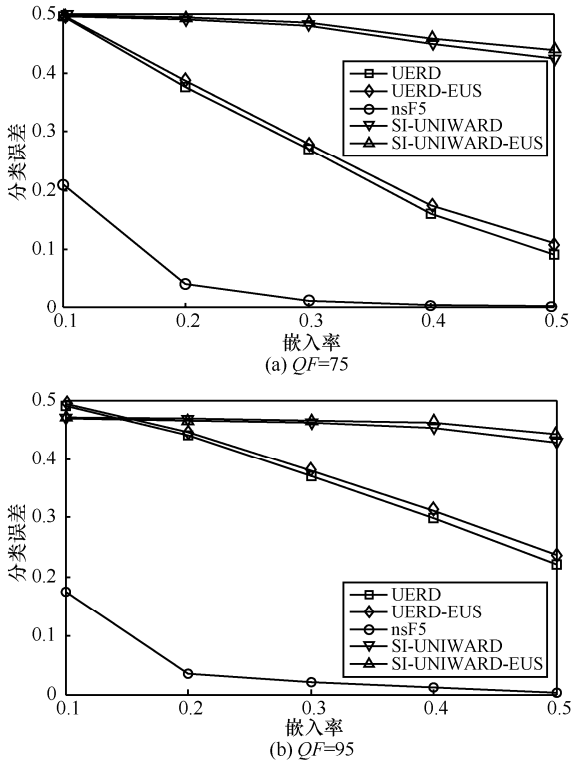


图 8　CC-JRM 特征在不同质量因子下的检测结果

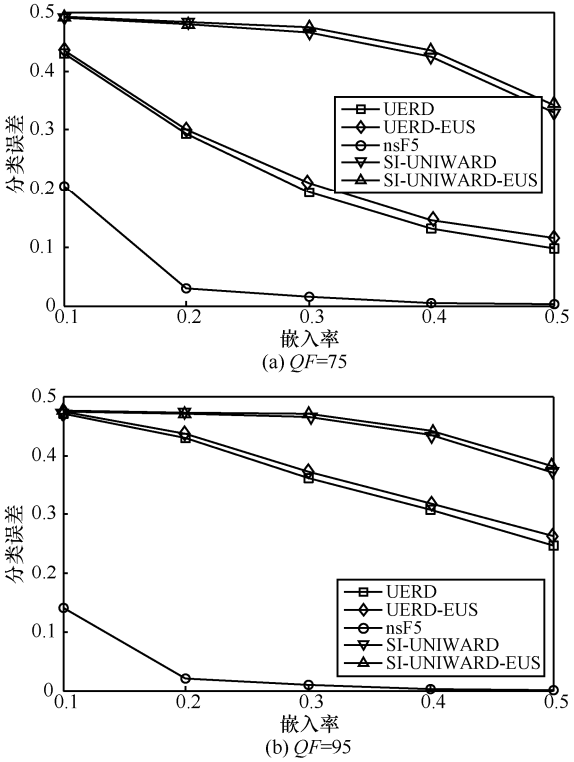


图 9　DCTR 特征在不同质量因子下的检测结果

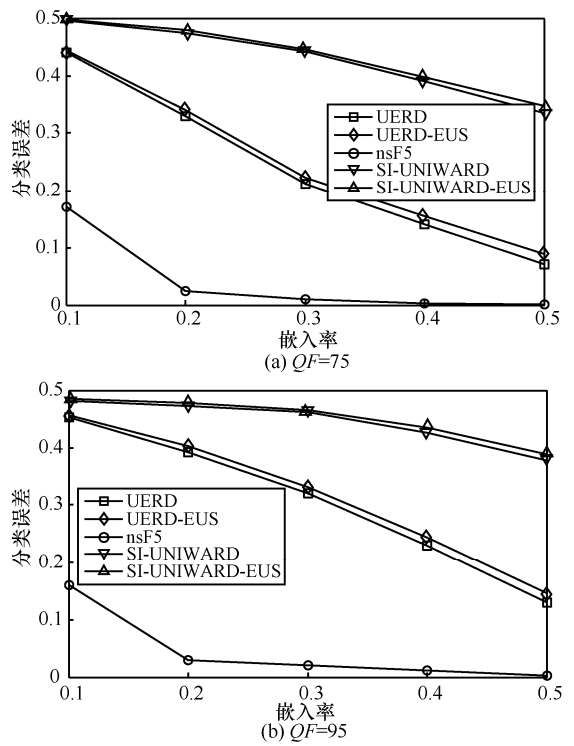


图 10　GFR 特征在不同质量因子下的检测结果

随着嵌入率的增加，UERD-EUS、SI-UNIWARD-EUS 算法和与之相对应的 UERD、SI-UNIWARD 算法相比，抗检测性能提高得越来越明显。原因在于随着嵌入率的提高，EUS 策略能更好地保持 JPEG

图像的相关性，因此，提高了算法的安全性。此外，由于 nsF5 是非自适应隐写算法，安全性较弱，因此，在 2 种隐写检测特征下的抗检测性能普遍偏低。

6 结束语

本文提出一种失真代价更新策略 EUS，使失真代价能够分批次地动态更新，并结合该更新策略设计并实现了一种自适应 JPEG 隐写算法。实验表明，该算法能够有效保持 DCT 系数相关性，并能显著提高 JPEG 隐写的安全性。在未来的工作中考虑用更加精确的衡量尺度计算嵌入波动，并用更加科学的方式分配秘密信息比特，进一步完善算法性能。

参考文献：

[1] SEDIGHI V, COGRANNE R, FRIDRICH J. Content-adaptive steganography by minimizing statistical detectability[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(2): 221-234.

[2] FILLER T, FRIDRICH J. Design of adaptive steganographic schemes for digital images[C]//IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics. 2011: 78800F-78800F-14.

[3] FRIDRICH J, GOLJAN M, SOUKAL D. Perturbed quantization steganography with wet paper codes[C]//The 2004 Workshop on Multimedia and Security. 2004: 4-15.

[4] KIM Y, DURIC Z, RICHARDS D. Modified matrix encoding technique for minimal distortion steganography[C]//International Workshop on Information Hiding. Springer Berlin Heidelberg. 2006: 314-327.

[5] FRIDRICH J, PEVNÝ T, KODOVSKÝ J. Statistically undetectable jpeg steganography: dead ends challenges, and opportunities[C]//The 9th Workshop on Multimedia & Security. 2007: 3-14.

[6] FILLER T, JUDAS J, FRIDRICH J. Minimizing additive distortion in steganography using syndrome-trellis codes[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3): 920-935.

[7] WANG C, NI J. An efficient JPEG steganographic scheme based on the block entropy of DCT coefficients[C]//2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2012: 1785-1788.

[8] GUO L, NI J, SHI Y Q. Uniform embedding for efficient JPEG steganography[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(5): 814-825.

[9] GUO L, NI J, SU W, et al. Using statistical image model for JPEG steganography: uniform embedding revisited[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(12): 2669-2680.

[10] HUANG F, LUO W, HUANG J, et al. Distortion function designing for JPEG steganography with uncompressed side-image[C]//The 1st ACM Workshop on Information Hiding And Multimedia Security. 2013: 69-76.

[11] HOLUB V, FRIDRICH J. Digital image steganography using universal distortion[C]//1st ACM Workshop on Information Hiding and Multimedia Security. Montpellier, France, 2013: 59-68.

[12] DENEMARK T, FRIDRICH J. Side-informed steganography with additive distortion[C]//IEEE International Workshop on Information Forensics and Security (WIFS), 2015: 1-6.

[13] LI B, WANG M, LI X, et al. A strategy of clustering modification directions in spatial image steganography[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(9): 1905-1917.

[14] BAS P, FILLER T, PEVNÝ T. Break our steganographic system: the ins and outs of organizing BOSS[C]//International Workshop on Information Hiding. 2011: 59-70.

[15] KODOVSKÝ J, FRIDRICH J. Steganalysis of JPEG images using rich models[C]//IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics. 2012: 83030A-83030A-13.

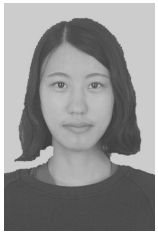
[16] HOLUB V, FRIDRICH J. Low-complexity features for JPEG steganalysis using undecimated DCT[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(2): 219-228.

[17] SONG X, LIU F, YANG C, et al. Steganalysis of adaptive JPEG steganography using 2D Gabor filters[C]//The 3rd ACM Workshop on Information Hiding and Multimedia Security. 2015: 15-23.

[18] KODOVSKY J, FRIDRICH J, HOLUB V. Ensemble classifiers for steganalysis of digital media[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2): 432-444.

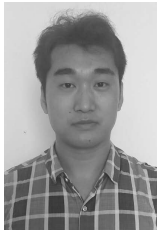
作者简介：

汤光明（1963-），女，湖南衡阳人，中国人民解放军信息工程大学教授、博士生导师，主要研究方向为信息安全、信息隐藏。



孙艺（1992-），女，吉林四平人，中国人民解放军信息工程大学硕士生，主要研究方向为信息隐藏。

徐潇雨（1993-），男，江苏连云港人，中国人民解放军信息工程大学硕士生，主要研究方向为深度学习。



王宇（1984-），男，吉林辽源人，吉林省四平军分区战备建设处参谋，主要研究方向为信息安全。