

图像隐写系统模糊相对熵安全性测度研究

欧阳春娟^{1,2}, 李霞^{1,2}, 李斌^{1,2}

(1. 深圳大学信息工程学院, 广东深圳 518060; 2. 深圳市现代通信与信息处理重点实验室, 广东深圳 518060)

摘要: 模糊相对熵可很好地度量两个模糊集之间的差异. 文章根据隐写通信过程的不确定性, 定义了隐写系统 n 阶 Markov 链模型的模糊经验矩阵, 提出了隐写系统的模糊相对熵和加权模糊相对熵安全性测度, 证明了该安全性测度的非负性、交换性和一致性. 此外, 由该测度可推导出各种确定模式下安全性测度. 仿真实验表明, 与同模型下的确定模式安全性测度相比, 模糊相对熵及加权模糊相对熵安全性测度对隐写算法安全性的度量能力更强, 且随着阶数的增加, 对应安全性测度的度量能力增强. 隐写算法设计实验也表明模糊相对熵安全性测度可更好地指导设计高安全性的隐写算法.

关键词: 隐写系统; 模糊相对熵; 马尔可夫模型; 安全性测度

中图分类号: TP309.2 **文献标识码:** A **文章编号:** 0372-2112 (2012) 08-1515-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2012.08.004

Fuzzy Relative Entropy Security Measures for Image Steganographic System

OUYANG Chun-juan^{1,2}, LI Xia^{1,2}, LI Bin^{1,2}

(1. College of Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China

2. Shenzhen Key Laboratory of Modern Communications and Information Processing, Shenzhen, Guangdong 518060, China)

Abstract: Fuzzy relative entropy is capable of measuring the difference between two fuzzy sets. According to the indeterminacy of steganography communication, a fuzzy empirical matrix for n -th order Markov model is defined. New security measures in terms of fuzzy relative entropy and weighted fuzzy relative entropy are introduced for steganographic system. These new security measures are proved to be nonnegative, symmetric and uniform. Furthermore, some existing security measures under a deterministic data statistical distribution model can be derived from the proposed security measures. Simulation results show the new security measures have better evaluating ability than the existing deterministic security measures under the same modeling condition. In addition, the higher the order of the Markov model, the better the measuring ability of the proposed security measures. The proposed security measures may also provide more insights for designs of secure steganographic algorithms.

Key words: steganographic system; fuzzy relative entropy; Markov model; security measure

1 引言

隐写术^[1]是信息隐藏技术的一个重要分支, 其研究如何将秘密信息隐藏于公开的数字媒体中而实现隐蔽通信. 和所有的保密通信技术一样, 安全性是其首要的研究问题. 然而, 相对于隐写和隐写分析算法的研究而言, 隐写系统安全性理论研究进展较为缓慢. 正如 Ross Anderson^[2]所言, 信息安全的理论研究在某种程度上比技术研究更为重要. 目前, 隐写及隐写分析技术^[3,4]发展迅速, 因此, 研究具有指导设计高安全隐写及高性能

隐写分析算法的隐写系统安全性评估方法非常必要.

对于隐写系统安全性评价指标而言, 其取决于载体图像与载密图像之间的统计分布差异, 而与具体的隐写及隐写分析算法无关. 所以建立合适的图像分布模型成为定义隐写系统安全性的关键之一. Cachin 等^[5]将隐写建模为统计中的假设检验问题. 假设载体数据与载密数据的分布为两个随机变量, 以这两个分布之间的相对熵 (也称为 Kullback-Leibler, KL 距离) 定义了隐写系统的安全性. Sullivan 等^[6]将图像构建成 1 阶 Markov 链模型, 考虑每个像素只与其相邻一个像素相关, 根据该模型经验

矩阵的散度距离定义隐写系统安全性.为了更好地刻画像素相关性,张湛等^[7,8]将图像扫描序列构建成 n 阶 Markov 模型,以该模型经验矩阵的散度距离提出隐写系统安全性测度.该模型较全面地考虑了图像像素相关性,通过对模型阶数的调节,可调整安全性统计测度的计算复杂度.Pevný 等^[9]认为图像在高维空间计算 KL 距离非常困难,但高阶特征可以达到很好的隐写分析效果,其利用两个样本统计的最大均差异 MMD(Maximum Mean Discrepancy)来计算载体与载密数据高维特征的差异,定义了隐写系统的 MMD 安全性测度.

信息论的创始人 Shannon 指出^[10],通信的作用就是通过消息的传递,使接收者从收到的消息中获取一定的信息,从而消除原来的不确定性.Shannon 从“不确定性”观点出发,给“信息”下了明确的定义.隐写作为一种信息隐藏通信技术,在通信过程中,图像各个局部细节都会发生不确定性变化,导致其各种特征统计变化也是不确定的.此外,在实际隐写通信过程中,一般无法获得载体数据,也无法获得无限的载密数据样本.因此,就无法获得载体数据和载密数据确定的统计分布.以上文献都从载体数据确定的分布模式下讨论安全性,不能有效地描述隐写通信的不确定性.为此,本文将隐写通信过程建模为模糊不确定性过程,以 n 阶 Markov 链作为图像统计分布模型,定义该模型的模糊转移经验矩阵.提出了隐写系统的模糊相对熵及加权模糊相对熵安全性测度.并证明了该测度的非负性,交换性和一致性.当模糊经验矩阵中元素完全隶属于对应的模糊集时,根据不同阶 Markov 链模型可推导出各种确定模式下安全性测度.实验表明,与相同模型确定模式下的安全性测度相比,模糊相对熵和加权模糊相对熵安全性测度可更有效度量不同隐写算法在不同嵌入率下的安全性,对隐写算法设计具有更好的指导作用.实验还表明加权模糊相对熵安全性测度比模糊相对熵安全性测度具有更强的度量能力.同时,随着 Markov 链模型阶数的增加,相对应阶数的安全性测度量度能力增强.

2 图像隐写系统数据的模糊经验矩阵

将图像扫描序列定义为 n 阶 Markov 链可更全面地描述像素之间的相关性,适合作为图像像素的统计模型^[7,8].本节通过定义隐写系统中载体数据与载密数据 n 阶 Markov 链的模糊经验矩阵来描述图像隐写引起的像素转移不确定性.

定义 1 (图像扫描序列 n 阶马尔可夫链) 对灰度图像按某种扫描方式(如按行,按列,按 zig-zag,按 Hilbert 等方式)得到序列 $X = \{x_1, x_2, \dots, x_t, \dots, x_L\}$.对序列中任意元素 $x_t, t \in [1, L]$ 只与前 $n-1$ 个元素相

关,即 $P(x_t | x_{t-1}, x_{t-2}, \dots, x_1) = P(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-n})$,称 X 为图像扫描序列的 n 阶 Markov 链,其中 L 为链长.

由定义 1 可得,当序列 X 中的像素与其他像素无关,图像像素的分布模型为独立同分布模型;当序列 X 中的元素只与前一个相邻像素相关,图像扫描序列为 1 阶 Markov 链模型.以此类推,可得到不同阶的图像扫描序列的 Markov 链模型.

定义 2 (n 阶 Markov 链的模糊经验矩阵) 设 X 为灰度图像扫描序列. $\eta_{i_1, i_2, \dots, i_{n+1}}(X)$ 为 X 中数据从 i_1 经过 i_2, i_3 等状态最终到达 i_{n+1} 的变换次数,即在 n 阶 Markov 链中序列 $i_1, i_2, \dots, i_n, i_{n+1}$ 出现的次数,则 $\eta_{i_1, i_2, \dots, i_{n+1}}(X)/(L-n)$ 为 n 阶 Markov 链 X (链长为 L) 中灰度像素值从 i_1 经过 i_2, i_3 等状态到达 i_{n+1} 的像素变换在总像素变换中占的比例,即对联合分布概率 $P(x_t = i_{n+1}, x_{t-1} = i_n, x_{t-2} = i_{n-1}, \dots, x_{t-(n+1)} = i_2, x_{t-n} = i_1)$ 的估计,定义, $M^X = \{m_{i_1, i_2, \dots, i_{n+1}} = \eta_{i_1, i_2, \dots, i_{n+1}}(X)/(L-n), i_k \in [0, 255]\}$ 为 n 阶 Markov 链 X 的经验矩阵, $m_{i_1, i_2, \dots, i_{n+1}}$ 为经验矩阵元素.其模糊经验矩阵定义为:

$$MF^X = \{\mu(m_{i_1, i_2, \dots, i_{n+1}}) | m_{i_1, i_2, \dots, i_{n+1}} \in M^X, i_k \in [0, 255]\} \quad (1)$$

其中 $\mu(m_{i_1, i_2, \dots, i_{n+1}})$ 为模糊经验矩阵元素,表示 $m_{i_1, i_2, \dots, i_{n+1}}$ 对模糊集 MF^X 的隶属度.

定义 3 (隐写系统数据的模糊经验矩阵) 设 C 和 S 分别为按某种扫描方式(如按行,按列,按 zig-zag,按 Hilbert 等方式)得到的隐写系统中载体数据集合和载密数据集合的 n 阶 Markov 链(链长为 L). $\eta_{i_1, i_2, \dots, i_{n+1}}(C)$ 和 $\eta_{i_1, i_2, \dots, i_{n+1}}(S)$ 为 C 和 S 中像素值从 i_1 经 i_2, i_3 等状态最终到达 i_{n+1} 的变换次数.则 C 和 S 的经验矩阵为:

$$M^C = \{m_{i_1, i_2, \dots, i_{n+1}} = \eta_{i_1, i_2, \dots, i_{n+1}}(C)/(L-n), i_k \in [0, 255]\} \quad (2)$$

$$M^S = \{m_{i_1, i_2, \dots, i_{n+1}} = \eta_{i_1, i_2, \dots, i_{n+1}}(S)/(L-n), i_k \in [0, 255]\} \quad (3)$$

$m_{i_1, i_2, \dots, i_{n+1}}$ 为经验矩阵 M^C 和 M^S 中元素, $m_{i_1, i_2, \dots, i_{n+1}}$ 所有可能的取值组成论域 $M_{i_1, i_2, \dots, i_{n+1}}$.则载体数据和载密数据 n 阶 Markov 链的模糊经验矩阵为:

$$MC = \{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) | m_{i_1, i_2, \dots, i_{n+1}} \in M_{i_1, i_2, \dots, i_{n+1}}, i_k \in [0, 255]\} \quad (4)$$

$$MS = \{\mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}) | m_{i_1, i_2, \dots, i_{n+1}} \in M_{i_1, i_2, \dots, i_{n+1}}, i_k \in [0, 255]\} \quad (5)$$

其中 $\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})$ 和 $\mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}})$ 表示元素 $m_{i_1, i_2, \dots, i_{n+1}}$ 对模糊集 MC 和 MS 的隶属度.

性质 1 MC 和 MS 为隐写系统中载体数据与载密数据 n 阶 Markov 链的模糊经验矩阵,其元素满足:若 $m_i > m_j$, 则 $\mu_{MC}(m_i) > \mu_{MC}(m_j)$, $\mu_{MS}(m_i) > \mu_{MS}(m_j)$, 且 $\sum_{i=1}^N \mu_{MC}(m_i) = 1$, $\sum_{i=1}^N \mu_{MS}(m_i) = 1$, N 为模糊经验矩阵 MC 和 MS 元素个数.

3 图像隐写系统模糊相对熵安全性

对于隐写系统,安全性是其最重要的评价指标.模糊相对熵可很好地度量两模糊集之间的差异^[11].本节根据隐写过程的不确定性定义图像隐写系统的模糊相对熵及加权模糊相对熵安全性测度.通过调节 n 阶 Markov 链模型阶数可控制安全性测度计算复杂度并得到不同阶的安全性测度.

定义 4 (隐写系统模糊相对熵安全性测度) 对于隐写系统, C 和 S 分别为载体数据样本集合和载密数据样本集合的 n 阶 Markov 链, M^C 和 M^S 分别为 C 和 S 的经验矩阵, MC 和 MS 分别 M^C 和 M^S 对应的模糊经验矩阵. $m_{i_1, i_2, \dots, i_{n+1}}$ 代表了 C 和 S 中像素值从 i_1 经 i_2, i_3 等状态到达 i_{n+1} 的模糊联合概率分布. G 为 i_1, i_2, \dots, i_{n+1} 所有可能的取值集合. $\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})$ 和 $\mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}})$ 分别为 MC 和 MS 中元素, 表示经验矩阵 M^C 和 M^S 中元素 $m_{i_1, i_2, \dots, i_{n+1}}$ 属于模糊集 MC 和 MS 的隶属度. 则隐写系统的模糊相对熵安全性测度定义为:

$$H_n(MC, MS) = \sum_{i_1, i_2, \dots, i_{n+1} \in G} [\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) \cdot \log \frac{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}{(\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2} + (1 - \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})) \cdot \log \frac{(1 - \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}))}{1 - (\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2}] \quad (6)$$

为使该模糊相对熵具有对称性,通常采用以下形式来定义隐写系统的模糊相对熵安全性测度.

$$D_n(MC, MS) = H_n(MC, MS) + H_n(MS, MC) \quad (7)$$

当 $D_n(MC, MS) = 0$, 称此隐写系统绝对安全; 当 $0 \leq D_n(MC, MS) \leq \epsilon$, 称此隐写系统 ϵ -安全.

定理 1 设 $D_n(MC, MS)$ 为图像隐写系统模糊相对熵安全性测度, 则 $D_n(MC, MS)$ 具有以下性质:

- (1) 非负性: $D_n(MC, MS) \geq 0$;
- (2) 对称性: $D_n(MC, MS) = D_n(MS, MC)$;
- (3) 一致性: $D_n(MC, MS) = 0$ 当且仅当 $MC = MS$.

证明:(1)非负性

$$- \sum_{i_1, i_2, \dots, i_{n+1} \in G} \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) \cdot \log \frac{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}{(\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2}$$

$$= \sum_{i_1, i_2, \dots, i_{n+1} \in G} \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) \cdot \log \frac{(\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2}{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}$$

由于 $f(x) = \log x$ 是凸函数, 由 Jensen 不等式, 上式可得:

$$\leq \log \sum_{i_1, i_2, \dots, i_{n+1} \in G} \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) \cdot \frac{(\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2}{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}$$

$$= \log \sum_{i_1, i_2, \dots, i_{n+1} \in G} (\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2$$

$$= \log(\frac{1}{2} \sum_{i_1, i_2, \dots, i_{n+1} \in G} \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \frac{1}{2} \sum_{i_1, i_2, \dots, i_{n+1} \in G} \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))$$

由性质 1, 上式可得:

$$= \log(\frac{1}{2} + \frac{1}{2}) = \log 1 = 0$$

取 $\mu'_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) = 1 - \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})$; $\mu'_{MS}(m_{i_1, i_2, \dots, i_{n+1}}) = 1 - \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}})$, 则同理可证 $-\sum_{i_1, i_2, \dots, i_{n+1} \in G} (1 - \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})) \cdot \log \frac{1 - \mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}{1 - (\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2} \leq 0$

因此, $-H_n(MC, MS) \leq 0$ 即 $H_n(MC, MS) \geq 0$.

同理可证 $H_n(MS, MC) \geq 0$, 所以, $D_n(MC, MS) \geq 0$

(2)对称性

$D_n(MC, MS) = H_n(MS, MC) + H_n(MC, MS)$, 且 $D_n(MS, MC) = H_n(MS, MC) + H_n(MC, MS)$ 因此, $D_n(MC, MS) = D_n(MS, MC)$.

(3)一致性

在证明(1)中, 由于 $\log x$ 为凸函数, 因此, 不等式等号成立的充分必要条件为:

$\frac{\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}})}{(\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}}))/2}$ 处处为 1, 即 $\mu_{MC}(m_{i_1, i_2, \dots, i_{n+1}}), \mu_{MS}(m_{i_1, i_2, \dots, i_{n+1}})$ 处处相等, 即 MC 和 MS 处处相等. 因此, 当且仅当 $MC = MS$ 时 $D_n(MC, MS) = 0$. 证毕.

为突出模糊经验矩阵中像素转移变化大的元素对总体模糊相对熵的贡献度, 对该元素在模糊经验矩阵中可分配一个较大的权重, 反之, 分配一个较小的权重. 因此, 通过引入权重向量^[12]进一步定义隐写系统加权模糊相对熵安全性测度.

定义 5 (隐写系统加权模糊相对熵安全性测度) 设隐写系统的模糊相对熵安全性测度为: $H_n(MC, MS)$. 引入权重向量 $\lambda_{i_1, i_2, \dots, i_{n+1}}, (i_1, i_2, \dots, i_{n+1} \in G)$, 隐

写系统加权模糊相对熵安全性测度定义为:

$$HW_n(\mathbf{MC}, \mathbf{MS}) = \sum_{i_1, i_2, \dots, i_{n+1} \in G} \lambda_{i_1, i_2, \dots, i_{n+1}} \left[\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) \frac{\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}})}{(\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}})) / 2} + (1 - \mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}})) \frac{(1 - \mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}))}{(1 - (\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) + \mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}})) / 2)} \right] \quad (8)$$

其中 $\lambda_{i_1, i_2, \dots, i_{n+1}} =$

$$\frac{|\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) - \mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}})|}{\sum_{i_1, i_2, \dots, i_{n+1} \in G} |\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) - \mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}})|}$$

考虑对称性,采用以下形式为隐写系统的加权模糊相对熵安全性测度:

$$DW_n(\mathbf{MC}, \mathbf{MS}) = HW_n(\mathbf{MC}, \mathbf{MS}) + HW_n(\mathbf{MS}, \mathbf{MC}) \quad (9)$$

当 $DW_n(\mathbf{MC}, \mathbf{MS}) = 0$, 称此隐写系统绝对安全; 当 $0 \leq DW_n(\mathbf{MC}, \mathbf{MS}) \leq \epsilon$, 称此隐写系统 ϵ -安全.

定理 2 假设 $DW_n(\mathbf{MC}, \mathbf{MS})$ 为图像隐写系统基于 n 阶 Markov 链模型的加权模糊相对熵安全性测度, 则 $DW_n(\mathbf{MC}, \mathbf{MS})$ 具有非负性, 对称性和一致性. (证明同定理 1, 略)

隐写系统载体数据集合和载密数据集合 n 阶 Markov 链模型的模糊经验矩阵元素 $\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}})$ 和 $\mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}})$ 表示经验矩阵中元素 $m_{i_1, i_2, \dots, i_{n+1}}$ 对模糊集 \mathbf{MC} 和 \mathbf{MS} 的隶属度. 当元素 $m_{i_1, i_2, \dots, i_{n+1}}$ 完全隶属于模糊集时, 模糊经验矩阵即为经验矩阵.

注 1 当 $\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$ 且 $\mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$ 时, $\mathbf{MC} = \mathbf{M}^C$, $\mathbf{MS} = \mathbf{M}^S$, 当 n 阶 Markov 链中 $n = 0$, 图像满足独立同分布模型, $\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) = P_c(x)$, $\mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}}) = P_s(x)$, 其中 $P_c(x)$, $P_s(x)$ 为载体数据集合与载密数据集合的像素概率分布, x 为像素值, X 为所有的像素取值集合, $P_c(x)$ 与 $P_s(x)$ 的相对熵如式 (10) 为文献 [5] 当从图像一阶统计特征角度考虑时的安全性测度.

$$C_0(P_c, P_s) = \sum_{x \in X} P_c(x) \log \frac{P_c(x)}{P_s(x)} \quad (10)$$

注 2 当 $\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$ 且 $\mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$, $n = 1$ 时, 图像分布模型为 1 阶 Markov 链模型, $\mathbf{MC} = \mathbf{M}_y^c$, $\mathbf{MS} = \mathbf{M}_y^s$, 其中 \mathbf{M}_y^c , \mathbf{M}_y^s 为载体数据与载密数据 1 阶 Markov 链的经验矩阵, 则 \mathbf{M}_y^c , \mathbf{M}_y^s 的散度距离如式 (11) 为文献 [6] 所提出的隐写系统 2 阶统计分布安全性测度.

$$C_1(\mathbf{M}_y^c, \mathbf{M}_y^s) = \sum_{i, j \in X} \mathbf{M}_y^c \log \left(\frac{\mathbf{M}_y^c}{\sum_j \mathbf{M}_y^c} \frac{\sum_j \mathbf{M}_y^s}{\mathbf{M}_y^s} \right) \quad (11)$$

注 3 当 $\mu_{\mathbf{MC}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$ 且 $\mu_{\mathbf{MS}}(m_{i_1, i_2, \dots, i_{n+1}}) = m_{i_1, i_2, \dots, i_{n+1}}$ 时, $\mathbf{MC} = \mathbf{M}^C$, $\mathbf{MS} = \mathbf{M}^S$. 当图像扫描序列为 n 阶 Markov 模型时, \mathbf{M}^C 和 \mathbf{M}^S 的散度距离如式 (12) 为文献 [7] 提出的隐写系统安全性测度.

$$C_n(\mathbf{M}^C, \mathbf{M}^S) = \sum_{i_1, i_2, \dots, i_{n+1} \in G} \left\{ m_{i_1, i_2, \dots, i_{n+1}}^c \log \left[\frac{m_{i_1, i_2, \dots, i_{n+1}}^c}{\sum_{i_1} m_{i_1, i_2, \dots, i_{n+1}}^c} \times \left(\frac{m_{i_1, i_2, \dots, i_{n+1}}^s}{\sum_{i_1} m_{i_1, i_2, \dots, i_{n+1}}^s} \right)^{-1} \right] \right\} \quad (12)$$

本节将隐写系统中载体数据集合与载密数据集合建模为 n 阶 Markov 链模型, 定义了 n 阶 Markov 链模型的隐写系统模糊相对熵安全性测度, 进一步提出了加权模糊相对熵安全性测度. n 的取值越大, 图像的 n 阶 Markov 链模型包含的像素相关性信息也更多. 但随着 n 的增加, 其计算复杂度将按指数级增大. 因此, 隐写系统模糊相对熵安全性测度 $D_n(\mathbf{MC}, \mathbf{MS})$ 和加权模糊相对熵安全性测度 $DW_n(\mathbf{MC}, \mathbf{MS})$, 在实际应用中, 一般只采用较低阶的 Markov 链模型. 特别地, 当图像扫描序列中的像素与其相邻像素无关, 图像满足独立同分布模型, 称 $D_0(\mathbf{MC}, \mathbf{MS})$ 和 $DW_0(\mathbf{MC}, \mathbf{MS})$ 为 0 阶模糊相对熵和 0 阶加权模糊相对熵安全性测度, 它反映了载体数据 1 阶统计分布改变情况. 当图像扫描序列为 1 阶 Markov 链模型, 称 $D_1(\mathbf{MC}, \mathbf{MS})$ 和 $DW_1(\mathbf{MC}, \mathbf{MS})$ 为 1 阶模糊相对熵和 1 阶加权模糊相对熵安全性测度, 它反映了载体数据 2 阶统计分布改变情况. 当图像扫描序列为 2 阶 Markov 链模型, 称 $D_2(\mathbf{MC}, \mathbf{MS})$ 和 $DW_2(\mathbf{MC}, \mathbf{MS})$ 为 2 阶模糊相对熵和 2 阶加权模糊相对熵统计分布测度, 它反映了载体数据高阶统计分布改变情况. 本文提出的隐写系统安全性测度充分考虑隐写通信的不确定性因素, 其采用模糊相对熵及加权模糊相对熵可较好地度量了隐写引起的载体数据的统计分布改变. 从隐写和隐写分析两方面看, 该安全性测度对研究载体统计分布改变较小的隐写算法, 以及为隐写分析提供高阶统计特征, 均可提供指导作用.

4 仿真实验与讨论

4.1 安全性度量能力比较

实验采用本文提出的隐写系统模糊相对熵和加权模糊相对熵安全性测度, 及确定模式下文献 [5] 中当从图像一阶统计特征角度考虑时的相关熵安全性测度 (记为 $C_0(PC, PS)$), 文献 [6] 中的散度距离安全性测度 (记为 $C_1(\mathbf{MC}, \mathbf{MS})$) 及文献 [7] 中取 2 阶 Markov 模型时的安全性测度 (记为 $C_2(\mathbf{MC}, \mathbf{MS})$), 来度量 LSBM (Least Significant Bits Matching, LSBM) 隐写 [13] 在不同嵌入率下安全性. 通过比较不同安全性测度对载体数据统计分

布的改变反映灵敏度来说明本文提出的安全性测度比确定模式下的安全性测度具有更强的安全性度量能力. 实验当中, $D_n(\mathbf{MC}, \mathbf{MS})$ 及 $DW_n(\mathbf{MC}, \mathbf{MS})$ 采用升半正态型隶属度函数 $\mu(x) = 1 - e^{-kx^2} (x > 0)$, 其中 $k = 2.5$ 生成载体数据及载密数据的模糊经验矩阵, 其满足随着经验矩阵中元素取值增大, 对应隶属度值增大的性质. 由文献[7, 14]可知, 对图像按 Hilbert 扫描所得的 Markov 链包含的像素相关性远多于按行(列)和按 zig-zag 扫描方式, 因此, 实验中采用 Hilbert 扫描图像得到对应的 Markov 链模型. 随机选择 NRCS 图像库^[15]中 500 幅图像, 裁剪为 512×512 , 并转化为灰度图像, 采用 LSBM 对图像进行隐写, 嵌入率从 0.05bpp(bits per pixel, bpp)增至 1bpp, 每次递增 0.05bpp 得到载密图像.

图 1 为不同安全性测度在同一嵌入率下的均值随嵌入率增加的变化曲线图, 其中取值为各测度均值的

对数. 其中图 1(a)为图像独立同分布模型下的安全性测度 $C_0(PC, PS)$, $D_0(\mathbf{MC}, \mathbf{MS})$ 和 $DW_0(\mathbf{MC}, \mathbf{MS})$ 针对 LSBM 在不同嵌入率下的针对 500 幅图的均值变化曲线图. 图 1(b)为 1 阶 Markov 链模型下的安全性测度 $C_1(\mathbf{MC}, \mathbf{MS})$, $D_1(\mathbf{MC}, \mathbf{MS})$, 及 $DW_1(\mathbf{MC}, \mathbf{MS})$ 针对 LSBM 在不同嵌入率下的针对 500 幅图的均值变化曲线图. 图 1(c)为 2 阶 Markov 链模型下的安全性测度 $C_2(\mathbf{MC}, \mathbf{MS})$, $D_2(\mathbf{MC}, \mathbf{MS})$ 和 $DW_2(\mathbf{MC}, \mathbf{MS})$ 针对 LSBM 在不同嵌入率下的针对 500 幅图的均值变化曲线图. 由图 1 可知, 每条曲线都满足随着嵌入率增加, 安全性测度取值增大, 安全性下降的规律. 同时图 1 显示在同种图像像素分布模型下, 加权模糊相对熵安全性的均值曲线斜率最大, 对载体数据统计分布变化反映最灵敏, 其次是模糊相对熵安全性的均值曲线.

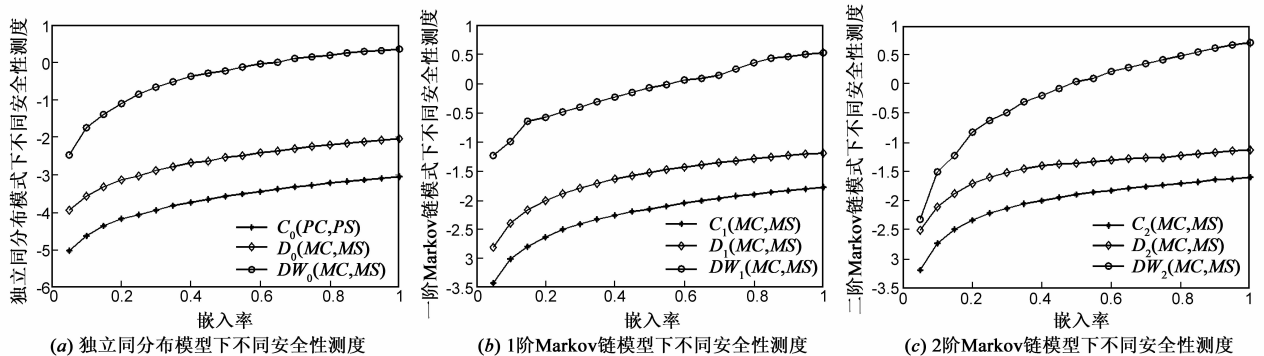


图 1 不同安全性测度对LSBM隐写在不同嵌入率下的安全性度量变化曲线图

为进一步验证模糊相对熵和加权模糊相对熵安全性测度对不同隐写算法在不同图像库的安全性度量的有效性. 采用以上实验中的各种安全性测度来度量 F3,

F4, F5 三种频率域隐写在不同嵌入率下的安全性, 嵌入率从 0.2bpac(bits per non-zero AC coefficient)增至 1bpac, 每次递增 0.2bpac. 选择图像库 BOWS2^[16]中 3000 幅大小为 512×512 的灰度图像进行实验.

表 1 不同安全性测度对不同隐写算法在不同嵌入率下的安全性度量均值

安全性 嵌入率	不同模型不同阶数的安全性测度									
	独立同分布模型			一阶 Markov 链模型			二阶 Markov 链模型			
	C_0	D_0	DW_0	C_1	D_1	DW_1	C_2	D_2	DW_2	
F3	0.2	-5.822	-5.689	-3.734	-4.246	-4.085	-3.043	-3.873	-3.746	-2.558
	0.4	-5.156	-5.072	-3.258	-3.772	-3.675	-2.587	-3.434	-3.374	-2.127
	0.6	-4.691	-4.383	-2.863	-3.446	-3.329	-2.378	-3.217	-3.128	-1.983
	0.8	-4.457	-4.281	-2.634	-3.286	-3.184	-2.125	-3.105	-2.989	-1.738
	1.0	-4.269	-4.137	-2.437	-3.019	-2.983	-1.962	-2.899	-2.736	-1.582
F4	0.2	-6.217	-6.023	-4.785	-5.943	-5.612	-4.389	-5.562	-5.373	-3.893
	0.4	-5.369	-5.217	-4.376	-5.212	-5.174	-4.171	-5.119	-5.079	-3.492
	0.6	-4.978	-4.863	-4.173	-4.743	-4.731	-4.004	-4.765	-4.699	-3.247
	0.8	-4.735	-4.661	-3.862	-4.587	-4.516	-3.791	-4.376	-4.442	-3.029
	1.0	-4.591	-4.437	-3.685	-4.475	-4.392	-3.572	-4.207	-4.282	-2.873
F5	0.2	-6.532	-6.347	-4.942	-6.012	-5.973	-4.523	-5.794	-5.532	-3.975
	0.4	-6.178	-6.013	-4.481	-5.782	-5.569	-4.312	-5.439	-5.361	-3.628
	0.6	-5.932	-5.891	-4.276	-5.573	-5.378	-4.213	-5.291	-5.116	-3.429
	0.8	-5.844	-5.632	-4.189	-5.391	-5.237	-4.038	-5.156	-5.003	-3.356
	1.0	-5.682	-5.528	-4.007	-5.284	-5.187	-3.899	-4.843	-4.892	-3.284

表 1 为不同安全性测度在不同嵌入率下对 3000 幅图像度量安全性的均值, 表中取值为各测度均值的对数值. 由表 1 可见, 针对同一种算法, 在不同嵌入率, 同模型下的三种不同的安全性测度的取值都是随着嵌入率的增加而增大, 即三种安全性测度均能度量隐写算法的安全性. 对不同的隐写算法, 采用同一安全性测度进行度量, 可得

出其安全度量取值按隐写 F3, F4, F5 的顺序降低,说明这三种隐写算法中的安全 F3 安全性最低, F5 安全性最高,该结论与三种隐写安全性理论分析相符合.表 1 数据还表明,针对表中任意一种隐写算法在同一嵌入率下,均可得到加权模糊相对熵安全性测度的度量值大于模糊相对熵安全性测度的取值;模糊相对熵安全性测度的取值大于同模型下相对熵的安全性测度取值,说明加权模糊相对熵安全性测度对隐写引起的载体数据统计特征变化最灵敏.

以上实验结果表明,模糊相对熵和加权模糊相对熵安全性测度对不同隐写算法在不同图像库的安全性度量具有通用性.与同模型确定模式下的安全性测度相比,本文提出的两种安全性测度具有更强的安全性度量能力.此外,加权模糊相对熵安全性测度比模糊相对熵安全性测度对隐写算法的安全性度量能力更强.

表 2 为针对图 1 中的不同安全性测度对 LSBM 隐写在嵌入率从 0bpp 到 1bpp 之间变化时安全性测度变化范围.由表 2 可得,对于三种不同阶的模糊相对熵和加权模糊相对熵安全性测度,阶数越高,其安全性测度变化总量越大,对隐写引起载体统计分布改变的反映更为灵敏,则其度量隐写系统安全性能力越强.由此可得,对于 $D_n(\mathbf{MC}, \mathbf{MS})$ 和 $DW_n(\mathbf{MC}, \mathbf{MS})$,随着阶数 n 增加,其对隐写引起的载体数据统计分布变化反映更充分,对应的安全性测度度量能力更强.从理论分析原因,是由于 n 越大,图像扫描序列的 n 阶 Markov 链包含的像素相关性越丰富,因此,其对应的安全性测度反映的载体分布变化越多.

表 2 不同安全性测度在嵌入率从 0 到 1 之间的安全性变化范围

安全性测度	变化范围	安全性测度	变化范围	安全性测度	变化范围
$DW_0(\mathbf{MC}, \mathbf{MS})$	2.2281	$D_0(\mathbf{MC}, \mathbf{MS})$	0.0113	$C_0(\mathbf{PC}, \mathbf{PS})$	0.0011
$DW_1(\mathbf{MC}, \mathbf{MS})$	3.3578	$D_1(\mathbf{MC}, \mathbf{MS})$	0.0651	$C_1(\mathbf{MC}, \mathbf{MS})$	0.0167
$DW_2(\mathbf{MC}, \mathbf{MS})$	5.029	$D_2(\mathbf{MC}, \mathbf{MS})$	0.0725	$C_2(\mathbf{MC}, \mathbf{MS})$	0.0251

4.2 安全性测度指导设计隐写算法能力比较

本实验通过比较采用模糊相对熵安全性测度

$D_n(\mathbf{MC}, \mathbf{MS})$ 与确定模式下隐写系统安全性测度,指导设计的隐写算法的抵抗隐写分析能力,来说明本文提出的安全性测度对隐写算法设计具有更好的指导作用.为了在同样条件下进行比较,对同样图像统计分布模型下安全性测度进行对比实验.分别采用 $D_0(\mathbf{MC}, \mathbf{MS})$ 与文献[5]当假设图像独立同分布下安全性测度 $C_0(\mathbf{PC}, \mathbf{PS})$; $D_1(\mathbf{MC}, \mathbf{MS})$ 与文献[6]在 1 阶 Markov 链模型的散度距离安全性测度 $C_1(\mathbf{MC}, \mathbf{MS})$; $D_2(\mathbf{MC}, \mathbf{MS})$ 与文献[7]中取 2 阶 Markov 链的散度距离安全性测度 $C_2(\mathbf{MC}, \mathbf{MS})$ 来指导设计隐写算法.采用粒子群算法优化 LSBM 隐写嵌入过程中需要修改的像素的加减 1 序列,通过保持图像的统计特征来提高隐写算法的安全性.在优化过程中,分别用以上六种安全性测度为优化目标,目标函数取值越小,隐写图像的安全性越好,得到以不同安全性测度为指导的改进 LSBM 隐写算法.例如以 $D_0(\mathbf{MC}, \mathbf{MS})$ 为优化目标记为 PSO-LSBD0 隐写,以 $C_0(\mathbf{PC}, \mathbf{PS})$ 作为优化目标记为 PSO-LSBC0 隐写.采用 NRCS 图像库中的 1542 幅图像,转为灰度图像,裁剪为 512×512 ,粒子数为 30,迭代次数为 20,用 PSO-LSBD0 和 PSO-LSBC0 隐写算法在嵌入率为 1bpp 下得到载密图像.采用文献[17]对以上二种隐写算法得到的载密图像及原图提取特征进行隐写分析,该文献基于小波系数提取了 78 维特征进行盲隐写分析,用 Fisher 作为分类器,其中 400 幅图像作为训练集,其余的为测试集,得到的 ROC(Receiver Operating Characteristic Curve, ROC)曲线图如图 2(a)所示.由图 2(a)可知,PSO-LSBD0 隐写获得比 PSO-LSBC0 隐写算法更小的 AUC(Area under ROC Curve, AUC)值,因此 PSO-LSBD0 隐写具有更强的抵抗隐写分析能力.图 2(b)为以安全性测度 $D_1(\mathbf{MC}, \mathbf{MS})$ 和 $C_1(\mathbf{MC}, \mathbf{MS})$ 为优化目标得到的改进隐写算法分别记为 PSO-LSBD1 和 PSO-LSBC1,按以上实验设计得到 ROC 曲线图.图 2(c)为以 $D_2(\mathbf{MC}, \mathbf{MS})$ 和 $C_2(\mathbf{MC}, \mathbf{MS})$ 为优化目标得到的改进隐写算法,分别记为 PSO-LSBD2 和 PSO-LSBC2,按以上实验设计得到 ROC 曲线图.由图 2(b)(c)可知,PSO-LSBD1 和 PSO-LSBD2 都取得了更低

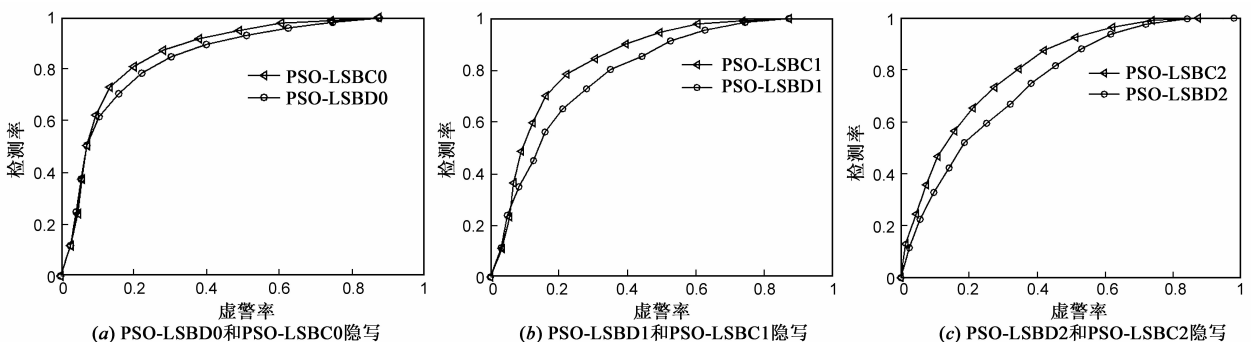


图 2 模糊相对安全性测度和确定模式下安全性测度指导设计的隐写算法 ROC 曲线图

的 AUC 值.此外,采用加权模糊相对熵进行同样的实验设置,可得到类似的结果.由此可得,模糊相对熵和加权模糊相对熵安全性测度与确定模式下的安全性测度相比,指导设计高安全性的隐写算法能力更强.

5 结论

根据隐写通信过程存在的各种不确定性因素,本文将图像扫描序列建模为 n 阶 Markov 链来描述像素相关性,定义隐写系统中载体数据和载密数据集合的 n 阶 Markov 链模型的模糊经验矩阵并讨论其性质.提出了隐写系统的模糊相对熵和加权模糊相对熵安全性测度.证明了这两种安全性测度的可行性.通过对 n 阶 Markov 链模型阶数调节,可得到不同阶隐写系统模糊相对熵和加权模糊相对熵安全性测度.当经验矩阵中元素完全隶属于该矩阵集合时,模糊经验矩阵等价于经验矩阵,则由模糊相对熵安全性测度可推导出各种确定模式下的隐写系统安全性测度.实验表明,模糊相对熵和加权模糊相对熵安全性测度均能有效地度量不同隐写算法在不同图像库及在不同嵌入率下的安全性.且随着阶数增加,其对应的 Markov 链模型包含图像相关性信息也相应增加,对隐写引起的图像统计分布改变反映更灵敏,相对应的安全性测度对隐写安全性的度量能力增强.与同模型确定模式下的安全性测度相比,本文提出的安全性测度具有更强的度量能力,其中加权模糊相对熵安全性测度比模糊相对熵安全性测度度量能力更强.此外,隐写算法设计实验也表明本文提出的安全性测度对隐写算法设计具有更强的指导能力.

n 阶马尔可夫链模型可较全面地描述数字图像相关性信息,通过对 n 的选择可灵活地调节模型的计算复杂度.因此,在实际应用中,通过分析隐写对其经验矩阵分布的改变,设计以模糊相对熵安全性测度为指导的隐写分析算法及高安全性的隐写算法均可作为下一步研究方向.

参考文献

- [1] 王朔中,张新鹏,张卫明.以数字图像为载体的隐写分析研究进展[J].计算机学报,2009,32(7):1247-1263.
Wang Shuo-Zhong, Zhang Xin-Peng, Zhang Wei-Ming. Recent advances in image-based steganalysis research [J]. Chinese Journal of Computers, 2009, 32(7): 1247-1263. (in Chinese)
- [2] Ross Anderson. Why information security is hard—an economic perspective [A]. Proceedings of 17th Annual Computer Security Applications Conference [C]. Washington, DC: IEEE Press, 2001. 39-40.
- [3] 张良.一种基于幅度预测的隐写分析方法[J].电子学报, 2010, 38(11): 2704-2707.
- [4] ZHANG Liang. A Steganalysis scheme using magnitude prediction [J]. Acta Electronica Sinica, 2010, 38(11): 2704-2707. (in Chinese)
- [4] 毛家发,钮心忻,杨义先,时书剑.基于 JPEG 净图定量描述的隐写分析方法[J].电子学报, 2011, 39(8): 1907-1912.
Mao Jia-fa, Niu Xin-xin, YANG Yi-xian, Shi Shu-jia. Steganalysis method based on JPEG cover image quantitative describing [J]. Acta Electronica Sinica, 2011, 39(8): 1907-1912. (in Chinese)
- [5] Cachin C. An information-theoretic model for Steganography [J]. Information and Computation, 2004, 192(1): 41-56.
- [6] Sullivan K, Madhow U, Chandrasekaran S, et al. Steganalysis for Markov cover data with applications to images [J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 275-287.
- [7] 张湛,刘光杰,王俊文等.基于图像高阶 Markov 链模型的扩频隐写分析[J].电子学报, 2010, 38(11): 2578-2584.
Zhang Zhan, Liu Guangjie, Wang Junwen, et al. Steganalysis of spread spectrum image steganography based on high-order Markov chain model [J]. Acta Electronica Sinica, 2010, 38(11): 2578-2584. (in Chinese)
- [8] 张湛,瞿芳,刘光杰等.基于高阶 Markov 链模型的数字图像隐写安全性评估方法[J].信息与控制, 2010, 39(4): 455-461.
Zhang Zhan, Qu Fang, Liu Guangjie, et al. A novel security evaluation method for digital image steganography based on high-order Markov chain model [J]. Information and Control, 2010, 39(4): 455-461. (in Chinese)
- [9] Pevný T, J Fridrich. Benchmarking for steganography [A]. Proceedings of the 10th Information Hiding International Workshop [C]. Berlin: Springer Verlag, 2008. 251-267.
- [10] 王维琼.模糊信息度量的拓展及应用[D].西北大学, 2005. 7-10.
Wang Weiqiong. Expansion and application of the fuzzy information measure [D]. Northwest University, 2005. 7-10. (in Chinese)
- [11] Xie Weixin, Bedrosian S D. Information measure for fuzzy sets [J]. IEEE Transactions on System, Man and Cybernetics, 1984, 14(1): 151-156.
- [12] 胡为,胡静涛.加权模糊相对熵在电机转子故障模糊识别中的应用[J].信息与控制, 2009, 38(3): 326-329.
Hu Wei, Hu Jingtao. Application of weighted fuzzy relative entropy to fuzzy recognition of motor rotor fault [J]. Information and Control, 2009, 38(3): 326-329. (in Chinese)
- [13] Sharp T. An implementation of key-based digital signal steganography [J]. Proceedings of Information Hiding Workshop, 2001, 2137: 13-26.

- [14] 戴跃伟,刘光杰,叶曙光.基于 Hilbert 填充曲线的自适应隐写[J].电子学报,2008,36(12A):35-38.
Dai Yuewei, Liu Guangjie, Ye Shuguang. Adaptive steganography based on Hilbert filling curve [J]. Acta Electronica Sinica, 2008, 36(12A): 35-38. (in Chinese)
- [15] United States Department of Agriculture, Natural resources conservation service photo gallery [DB/OL]. <http://photo>

[gallery.nrcs.usda.gov](http://photo.gallery.nrcs.usda.gov), 2002.

- [16] BOWS-2 database of 10,000 watermarked images [DB/OL]. <http://bows2.gipsa-lab.inpg.fr/BOWS2Image DataBase.tgz>. 2008.
- [17] Xuan G, Shi Y, Gao J, et al. Steganalysis based on multiple features formed by statistical moments of wavelet characteristic functions [J]. Computer Science, 2005, 3727: 262-265.

作者简介



欧阳春娟 女, 1974 年出生于江西省吉安市. 深圳大学信息工程学院博士研究生, 主要研究方向为信息隐藏及隐写分析.

E-mail: oycj001@163.com



李斌(通讯作者) 男, 1982 年出生于广东省深圳市. 深圳大学讲师、博士. 研究方向为信息隐藏及模式识别.

E-mail: libin@szu.edu.cn



李霞 女, 1968 年出生于四川省乐山市. 深圳大学教授、博士生导师. 主要研究方向为智能优化、智能计算及应用.

E-mail: lixia@szu.edu.cn