文章编号：2096-1618(2016)01-0070-06

基于训练集构造的图像通用盲检测算法改进

段怀锋[1]， 杨 榆[2]， 雷 敏[2]， 王惠华[1]
(1. 北京印刷学院信息工程学院，北京 102600；2. 北京邮电大学信息安全中心，北京 100876)

摘要：现有通用盲检测技术普遍存在泛化问题，导致检测器实用性大大下降。根据正交设计原则构建隐写率失匹配集合，隐写算法失匹配集合和图像源失匹配集合，分别分析检测SPAM分析算法和Rich Model分析算法在隐写率失匹配，隐写算法失匹配和图像源失匹配方面的检测率。并根据测试结果提出通过训练小隐写率图像集，训练多类隐写算法，图像预分类和改进IQM分析算法几种方案解决泛化问题，实验结果显示经过改进后隐写分析算法性能得到明显提升。

关 键 词：通用盲检测；泛化能力；SPAM；Rich Model；实用性
中图分类号：TP309.2 **文献标志码**：A

0 引言

随着互联网和计算机技术的迅速发展，信息的篡改以及破坏使人们意识到信息安全问题日益突出。信息隐藏技术是信息安全领域的研究热点之一，隐写分析和隐写术是隐藏技术的2个重要分支，两者对立。隐写分析是对隐写算法进行分析检测以判断载体中是否隐藏秘密信息[1]，隐写分析的不断发展又推动着图像信息隐藏和水印算法的不断改进[2-3]。

目前主流的隐写分析方法主要采用机器学习的方法对秘密信息的存在性进行检测，先提取图像特征，对图像特征进行分析然后训练分类器进行检测分类。随着特征提取的有效性和分类器性能的提高，通用盲检测方法性能越来越好，并能检测多种算法[4]。现有隐写分析算法研究主要关注通用隐写分析特征的设计，但很少关注实用性。实际上，通用隐写分析算法存在与所有机器学习问题相同的泛化问题。泛化问题若得不到解决，现有通用隐写分析算法很难投入实际应用。

以国际上比较有代表性的2种隐写分析算法SPAM(subtractive pixel adjacency matrix)和Rich Model算法模型为基础，分析其存在的泛化问题，包括隐写率失匹配问题，隐写算法失匹配问题和图像源失匹配问题，并根据测试结果提出几种可行的改进方案提高通用盲检测技术的实用性。

1 算法分析

1.1 SPAM分析算法

Fridrich等[5]在2010年提出差分像素邻接模型SPAM分析方法，用一阶和二阶Markov链描述相邻像素之间的差值。SPAM同时考虑了8个方向的差值，将垂直正逆2个方向、水平正逆2个方向这4个方向差值的均值和主对角线正逆2个方向，副对角线正逆2个方向这4个方向差值的均值的统计直方图作为特征用于分类。

1.2 Rich Model分析算法

Fridrich等[6]在2012年针对信息隐藏的检测提出Rich Model空域隐写分析算法，算法从空域提取对隐写分析有用的各类别特征，形成106个特征子集，包含34671维有效特征，RM算法的基本步骤如下：

(1)计算残差：Rich Model算法首先通过高通滤波得到量化图像的噪声残差，组成噪声分量模型。输入一幅图像 X，通过式 $R_{ij}=\hat{X}_{ij}(N_{ij})-cX_{ij}$ 计算残差，c 是残差阶数，N_{ij} 是像素 X_{ij} 的相邻像素，$\hat{X}_{ij}$ 是 X_{ij} 的预测值。残差类型包括SPAM类型和MINMAX类型两类，SPAM残差可分为一阶(1st)、二阶(2nd)、三阶(3rd)、EDGE残差和SQUARE残差[7]，SQUARE残差仅3×3窗口大小为例，其他残差以一个方向为例，各计算公式如式(1)所示。

$$
\begin{aligned}
&1\text{st}: R_{ij}=X_{i,j+1}-X_{ij}\\
&2\text{nd}: R_{ij}=X_{i,j-1}+X_{i,j+1}-2X_{ij}\\
&3\text{rd}: R_{ij}=-X_{i,j+2}+3X_{i,j+1}-3X_{ij}+X_{i,j-1}\\
&\text{EDGE}: R_{ij}=-X_{i-1,j-1}+2X_{i-1,j}+X_{i-1,j+1}+2X_{i,j-1}-4X_{ij}\\
&\quad +2X_{i,j+1}\\
&\text{SQUARE}: R_{ij}=-X_{i-1,j-1}+2X_{i-1,j}+X_{i-1,j+1}+2X_{i,j-1}-\\
&\qquad 4X_{ij}+2X_{i,j+1}-X_{i+1,j-1}+2X_{i+1,j}-X_{i+1,j+1}
\end{aligned}
\tag{1}
$$

收稿日期：2016-02-02
基金项目：国家自然科学基金资助项目(61310306028)，浙江省自然科学基金资助项目(Y15F020053)

MINMAX 残差可分为一阶(1st)、二阶(2nd)和 EDGE 残差 3 种,其计算方式为

$$\begin{aligned}
&1st:\ R_{ij}=\min\text{ or }\max\{X_{i-1,j-1}-X_{ij},X_{i-1,j}-X_{ij},X_{i-1,j+1}-X_{ij},X_{i,j+1}-X_{ij}\}\\
&2nd:\ R_{ij}=\min\text{ or }\max\{X_{i,j-1}+X_{i,j+1}-2X_{ij},X_{i-1,j}+X_{i+1,j}-2X_{ij}\}\\
&EDGE:R_{ij}=\min\text{ or }\max\{-X_{i-1,j-1}+2X_{i-1,j}-X_{i-1,j+1}+2X_{i,j-1}-4X_{ij}+2X_{i,j+1},-X_{i-1,j+1}+2X_{i-1,j}-X_{i+1,j+1}+2X_{i,j+1}-4X_{ij}+2X_{i+1,j},-X_{i+1,j+1}+2X_{i,j+1}-X_{i+1,j-1}+2X_{i+1,j}-4X_{ij}+2X_{i,j-1}\}
\end{aligned}\tag{2}$$

(2)截断和量化:将得到的各残差进行截断和量化,得:

$$R_{ij}\leftarrow trunc_T(round(\frac{R_{ij}}{q}))\tag{3}$$

式中 $q>0$,为量化参数,合适的量化参数可以使残差对嵌入信息引起的变化更加敏感,T 是截断阈值,截断的目的是为了限制残差的在合适的动态范围之内,使共生矩阵可以更好的被计算。

(3)计算共生矩阵:在隐写分析技术中,主对角线和副对角线方向的共生矩阵对隐写分析的作用非常小,因此 Rich Model 分析算法只采用水平方向和垂直方向的共生矩阵 $C_d^{(h)}$ 和 $C_d^{(v)}$:

$$\begin{aligned}
&C_d^{(h)}=\frac{1}{Z}\mid\{R_{ij},R_{i,j+1},R_{i,j+2},R_{i,j+3}\mid R_{i,j+k-1}=dk,k=1,\ldots,4\}\mid\\
&C_d^{(v)}=\frac{1}{Z}\mid\{R_{ij},R_{i+1,j},R_{i+2,j},R_{i+3,j}\mid R_{i+k-1},j=dk,k=1,\ldots,4\}\mid
\end{aligned}\tag{4}$$

其中 Z 是确保 $\sum_{d\in T4}C_d^{(h)}=1$ 和 $\sum_{d\in T4}C_d^{(v)}=1$ 的归一化因子,截断阈值 $T=2$,对于每个共生矩阵有 $(2T+1)^4=625$ 维特征。

(4)由于量化和截断,以及共生矩阵本身具有的对称性,每个 SPAM 特征子集的共生矩阵特征由 2×625 维降到 338 维,降维原则为

$$\begin{aligned}
\bar{C}_d&\leftarrow C_d+C_{-d}\\
\bar{\bar{C}}_d&\leftarrow\bar{C}_d+\bar{C}_{\overleftarrow{d}}
\end{aligned}\tag{5}$$

其中 $d=(d_1,d_2,d_3,d_4)\in T_4$,$\overleftarrow{d}=(d_4,d_3,d_2,d_1)$,$-d=(-d_1,-d_2,-d_3,-d_4)$。

每个 MINMAX 特征子集的共生矩阵特征由 2×625 维降到 325 维,降维原则如式(6)所示。

$$\begin{aligned}
\bar{C}_d&\leftarrow C_d^{(\min)}+C_{-d}^{(\max)}\\
\bar{\bar{C}}_d&\leftarrow\bar{C}_d+\bar{C}_{\overleftarrow{d}}
\end{aligned}\tag{6}$$

其中 $C^{(\min)}$ 和 $C^{(\max)}$ 是相同残差下的“min”和“max”共生矩阵。

1.3　集成分类器

集成分类器[8-9]的思路是在对样本进行分类的时候,将相互之间具有独立决策能力的分类器联合起来组成一个强分类器,通常情况下集成分类器的预测能力要比单个分类器的预测能力好得多。采用随机森林[10]集成机器学习方法,它利用随机重采样技术 bootstrap 和节点随机分裂技术构建多棵决策树,通过投票得到最终分类结果。随机森林算法在当前数据集上相对其他算法有很大优势,它能处理高维数据并且具有较快的学习速度,同时在训练过程中能检测到各特征间的相互影响,容易做成并行化算法。

集成分类器的基分类器为 Fisher 线性分类器,Fisher 线性判别的基本原理就是要找到一个投影方向(线性变换,线性组合),将高维问题降低到一维解决,使 2 类样本在投影轴上的投影的交叠部分最少,达到最佳的分类效果。2 类样本投影的均值之差要尽可能的大,样本类内离散度要尽可能的小。

2　算法设计

实验所用图像库为 BOSSBase ver. 1.01 图像库[11]和地图图像库。BOSSBase 图像库由 10000 副分辨率为 512×512 的 pgm 格式的灰度图像组成,均为原始图像,没有经过其他额外处理,不会引入噪声,能较好地反映问题。地图图像库同样由 10000 副分辨率为 512×512 的 pgm 格式的灰度图像组成。用 8 种经典的隐写算法对原图嵌入水印作为水印图像样本库,每种算法图像样本库有 10000 张水印图,8 种经典算法分别为 LSB 隐写算法、LSB 匹配隐写算法[12]、边缘自适应(EA)隐写算法[13]、HUGO 隐写算法[14]、最优三元±1 编码(STC)隐写算法[15-16]、MLSB 隐写算法、SSTDM 隐写算法以及 PATCHWORK 隐写算法,以上 8 种隐写算法基本涵盖了图像空域所有的特征区域。实验对 SPAM 和 Rich Model 隐写算法的泛化性能进行测试,测试用的分类器为集成分类器(Ensemble Classifier),测试内容包括隐写率失匹配性能分析、算法失匹配性能分析和图像源失匹配性能分析。

2.1 隐写率失匹配性能分析

原图和水印图像 2 类样本特征之间的距离与隐写率有着重要的关系[17],因此首先对不同隐写率的水印图像进行测试。根据正交设计原则选取的水印图像的隐写率分别为10 %,50 % 和80 %。为提高实验结果的置信度,把 BOSSBase 图像库分为两组,9000 张水印图和 9000 张原图作为训练样本,1000 张不同隐写率的水印图和 1000 张原图作为测试样本,进行 10 次测试然后取平均值。Rich Model 分析算法的测试结果如表 1 所示,SPAM 分析算法的测试结果如表 2 所示。

从表 1 和表 2 可以看出在隐写率失匹配方面,就总体性能而言,Rich Model 分析算法检测率优于 SPAM 分析算法。就具体隐写算法而言,SPAM 分析算法对 LSB 算法,LSB 匹配算法和 MLSB 算法的检测错误率最低,对于另外几种算法的检测率一般。Rich Model 分析算法同样对 LSB 算法,LSB 匹配算法和 MLSB 算法具有较低的检测错误率,这和 LSB 类算法与原图的特征之间的距离较大有关,对于其他几种算法 Rich Model 同样具有较低的检测错误率。就不同隐写率而言,SPAM 分析算法和 Rich Model 分析算法情况一样,在训练样本隐写率为10 % 时,对隐写率为50 % 和 80 % 的测试样本检测错误率,对于大部分算法错误率都低于0.1。不过当训练样本隐写率为50 % 和80 % 时,对隐写率为10 % 的测试样本检测错误率较高,隐写率50 % 和80 % 两者相互测试结果差别不大,当训练样本隐写率为50 % 时略好于为80 % 时。

表 1 隐写率失匹配时 Rich Model 分析算法的检测错误率

隐写算法	隐写率					
	train 50 % test 10 %	train 10 % test 50 %	train 10 % test 80 %	train 80 % test 10 %	train 50 % test 80 %	train 80 % test 50 %
EA	0.5	0.0778	0.0772	0.4635	0	0.1148
STC	0.3227	0.0688	0.0602	0.4865	0.0195	0.224
LSB±1	0.4818	0.0198	0.016	0.4965	0.0055	0.0553
MLSB	0.5	0	0	0.5	0	0
SSTDM	0.394	0.0515	0.0423	0.4768	0.03	0.0948
LSB	0.5	0	0	0.5	0	0.0005

表 2 隐写率失匹配时 SPAM 分析算法的检测错误率

隐写算法	隐写率					
	train 50 % test 10 %	train 10 % test 50 %	train 10 % test 80 %	train 80 % test 10 %	train 50 % test 80 %	train 80 % test 50 %
EA	0.487	0.2835	0.44	0.496	0.067	0.131
STC	0.493	0.1145	0.1125	0.494	0.0385	0.2455
LSB±1	0.4795	0.0245	0.023	0.496	0.0065	0.0475
MLSB	0.5	0.0005	0.0005	0.5	0	0
SSTDM	0.4515	0.158	0.1485	0.4715	0.0875	0.1775
LSB	0.5	0	0	0.5	0	0

2.2 算法失匹配性能分析

隐写算法种类繁多,难以在训练阶段枚举所有算法。在实际应用中,检测器会碰到训练集中没有的隐写算法。同时不同隐写算法嵌入原理也不同,检测准确率变化较大,若此问题得不到解决,分析算法实用性会大大降低。针对此问题对 SPAM 分析算法和 Rich Model 分析算法进行了测试,为了提高实验结果的置信度,同样把 BOSSBase 图像库分为两组,9000 张水印图和 9000 张原图作为训练样本,1000 张和训练样本用不同的隐写算法的水印图和 1000 张原图作为测试样本,进行 10 次测试然后取平均值。Rich Model 分析算法的测试结果如表 3 所示,SPAM 分析算法测试结果如表 4 所示。

表 3　隐写算法匹配时 Rich Model 分析算法的检测错误率

测试隐写算法	训练隐写算法							
	EA	HUGO	STC	LSB±1	PATCH	MLSB	SSTDM	LSB
EA	0.023	0.0495	0.044	0.4255	0.4985	0.5	0.196	0.4385
HUGO	0.478	0.091	0.4445	0.5005	0.4995	0.5	0.4785	0.5
STC	0.4425	0.095	0.023	0.4985	0.5	0.5	0.4945	0.5
LSB±1	0.457	0.052	0.0165	0.0105	0.4985	0.5	0.237	0.5
PATCH	0.015	0.047	0.011	0.464	0.006	0.4955	0.253	0.4475
MLSB	0.2915	0.5325	0.362	0.195	0.3655	0	0.0985	0.0095
SSTDM	0.482	0.0615	0.4005	0.4985	0.5005	0.5	0.0255	0.5
LSB	0.4085	0.353	0.055	0.015	0.2605	0.0145	0.08	0

表 4　隐写算法失匹配时 SPAM 分析算法的检测错误率

测试隐写算法	训练隐写算法							
	EA	HUGO	STC	LSB±1	PATCH	MLSB	SSTDM	LSB
EA	0.103	0.4815	0.223	0.4465	0.303	0.4995	0.3905	0.4835
HUGO	0.5005	0.503	0.5	0.5	0.5	0.5	0.5	0.5
STC	0.461	0.511	0.076	0.5005	0.497	0.5	0.5065	0.5
LSB±1	0.5155	0.517	0.4905	0.011	0.4465	0.5	0.117	0.5
PATCH	0.0665	0.4835	0.044	0.5	0.025	0.5	0.3955	0.5
MLSB	0.205	0.501	0.523	0.006	0.509	0	0.0575	0.394
SSTDM	0.103	0.4815	0.223	0.4465	0.303	0.4995	0.3905	0.4835
LSB	0.5005	0.503	0.5	0.5	0.5	0.5	0.5	0.5

从表 3 和表 4 可以看出在隐写算法失匹配方面,就总体性能而言,Rich Model 分析算法检测率仍然要优于 SPAM 分析算法。就具体算法而言,当训练样本和测试样本所用隐写算法相同时检测错误率最低,当训练样本和测试样本所用隐写算法不同时性能差距较大,SPAM 分析算法和 Rich Model 分析算法对于嵌入原理差异不大的隐写算法有较好的检测能力(如 LSB 匹配算法对 LSB 算法,MLSB 算法对 LSB 算法,LSB 算法对 MLSB 算法等),SPAM 分析算法对 HUGO 隐写算法检测错误率特别高,为0.5005。Rich Model 分析算法覆盖的算法更多,检测错误率也更低,其中训练样本用 HUGO 算法时对大部分算法都有较低的检测错误率。

2.3　图像源失匹配性能分析

图像属性对隐写分析的准确率有较大影响,不同的图像源之间检测率会有很大的不同,当训练样本源和检测图像源不匹配时,检测的性能可能会大大降低[18]。匹配测试中把 BOSSBase 图像库分为两组,9000 张水印图和 9000 张原图作为训练样本,1000 张和训练样本用相同隐写算法嵌入的水印图和 1000 张原图作为测试样本。失匹配测试用的图像库为 BOSSBase 图像库和地图图像库,地图图像库和 BOSSBase 图像库特征差距较大,能更好地反映这一情况。从 BOSSBase 图像库中随机选取 9000 张水印图和 9000 张原图作为训练样本,从地图图像库随机选取 1000 张和训练样本用相同隐写算法嵌入的水印图和 1000 张原图作为测试样本。水印图像各算法的隐写率均为 50 %,进行 10 次测试然后取平均值。匹配测试各算法的检测错误率如表 5 所示,失匹配测试各算法的检测错误率如表 6 所示。

表 5　图像源匹配时 SPAM 分析算法和 Rich Model 分析算法的检测错误率

隐写分析算法	隐写算法						
	EA	MLSB	LSB	LSB±1	PATCH	SSTDM	STC
SPAM	0.105	0	0	0.0115	0.0255	0.1325	0.076
Rich Model	0.023	0	0	0.0105	0.006	0.0255	0.023

表 6　图像源失匹配时 SPAM 分析算法和 Rich Model 分析算法的检测错误率

隐写分析算法	隐写算法						
	EA	MLSB	LSB	LSB±1	PATCH	SSTDM	STC
SPAM	0.47	0.0715	0.019	0.2885	0.3155	0.494	0.4905
Rich Model	0.31	0.045	0.0615	0.4135	0.0055	0.324	0.07

从表5和表6可以看出在图像源失匹配时,就总体性能而言,SPAM分析算法和Rich Model分析算法性能都有不同程度的下降,但对于大部分隐写算法Rich Model总体性能还是要优于SPAM隐写算法。就具体算法而言,Rich Model分析算法对MLSB算法,LSB算法,PATCHWORK算法和STC算法仍有较低的检测错误率,SPAM分析算法对MLSB算法和LSB算法的检测错误率为0.0715和0.019,但对其他算法检测错误率都大于0.25。

3 研究改进

通过对实验数据的分析以及对各隐写算法原理的研究总结,提出几种改进通用盲检测技术泛化性能的方法:

(1)对隐写率失匹配问题,由于小隐写率训练样本对隐写率较大的训练样本有着较高的检测率,因此可以通过把小隐写率样本作为隐写分析的训练样本解决隐写率失匹配问题。

(2)对隐写算法失匹配问题,通过实验结果可以看到当隐写算法类型相同或嵌入原理差异不大时有着不错的检测率,因此使训练样本涵盖多种类型的隐写算法,尽量使训练样本包含现有的所有类型,通过这种方法可以提高隐写算法失匹配时的检测率。训练多类算法时Rich Model分析算法的检测错误率如表7所示,训练集所用隐写算法为HUGO算法,STC算法,SSTDM算法和LSB算法,几乎涵盖了现有空域算法的所有类型,可以看到对另外几种隐写算法均有较低的检测错误率,算法性能大大提升。

表7 训练多类算法时Rich Model分析算法的检测错误率

训练隐写算法	测试隐写算法			
	EA	MLSB	LSB±1	PATCH
HUGO STC SSTDM LSB	0.0335	0.02	0.033	0.0175

(3)对于图像源失匹配问题,可以通过提取检测图像特征对检测图像进行分类预处理,然后用对应的隐写分析器进行检测。也可以通过改进IQM分析算法,对检测图像进行滤波处理,根据滤波前后图像噪声差异大小进行判断是不是水印图,若差距较大则为水印图。

4 结束语

对SPAM和Rich Model算法原理进行介绍,并通过实验分析其在实际应用中的泛化性能,根据实验结果提出几种改进方案,通过训练小隐写率样本解决隐写率失匹配问题,通过训练多类型隐写算法解决隐写算法失匹配问题,以及通过对图像进行预分类和改进IQM分析算法解决图像源失匹配问题,提高了分析算法的泛化性能。未来可以对隐写分析算法对彩色图像的泛化性能进行测试,并且进行改进,提高其在彩色图像方面的实用性。

参考文献:

[1] Dumitrescu S, Wu X L, Wang Z. Detection of LSB Steganography via Sample Pair Analysis[J]. IEEE Trans. on Signal Processing, 2003, 51(7): 1995 – 2007.

[2] Zhihua Xia, Xinhui Wang, Xingming Sun, et al. Steganalysis of least significant bit matching using multi-order differences[J]. Security and Communication Networks, 2014, 7(8): 1283–1291.

[3] Zhihua Xia, Xinhui Wang, Xingming Sun, et al. Steganalysis of LSB matching using differences between nonadjacent pixels[C]. Multimedia Tools and Applications,2014.

[4] 万宝吉, 张涛, 侯晓丹,等. 基于Boosting算法融合的图像隐写分析方法[J]. 计算机工程, 2013, 39(12): 148–151.

[5] Pevny T, Bas P, Fridrich J. Steganalysis by SubtractivePixel AdjacencyMatrix[J]. IEEE Transactions onInformation Security and Forensics, 2010, 5(2): 215–224.

[6] Fridrich J, Kodovsky J. Rich Models for Steganalysis of Digital Images[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(3).

[7] 李彩霞. 空域自适应EA隐秘算法及分析算法的研究(D). 大连:大连理工大学, 2013.

[8] Kodovsky J, Fridrich J, Holub. V.: Ensemble Classifiers for Steganalysis of Digital Media. IEEE Trans[J]. Information Forensics Security, 2011.

[9] Kodovsky J, Pevny T, Fridrich J. Modern steganalysis can detect YASS[C]. In Proc. SPIE, Electronic Imaging, Security Forensics of MultimediaXII, San Jose, CA, 2010, 7541:201–211.

[10] Breiman L. Random forest[J]. Machine Learning, 2001, 45(1):5–32.

[11] Filler T, Pevny T, Bas P. BOSS (Break Our Steganography System) 2014 [EB/OL]. Availa-

ble: http://agents. fel. cvut. cz/stegodata/BossBase-1.01-cover. tar. bz2,2014.

[12] Sharp T. An Implementation of Key-based Digital Signal Steganography[C]. In Proceedings of the 4th International Workshop on Information Hiding, London, UK, 2001: 13-26.

[13] Luo W, Huang F, Huang J. Edge adaptive image steganographybased on LSB matching revisited [J]. IEEE Trans. Inform. Forensics Security, 2010, 5(2): 201 - 214.

[14] Pevny T, Filler T, Bas P. Eds.: Using High-dimensional Image Models to Perform Highly Undetectable Steganography[J], in Information Hiding, 12th Int. Workshop, Calgary, Canada, 2010, 6387: 161 - 177.

[15] Filler T, Fridrich J, Judas J. Minimizing Embedding Impact in Steganography Using Trellis-Coded Quantization [J]. In Proceedings SPIE, EI, Media Forensics and Security XII, San Jose, CA, 2010: 1-14.

[16] 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.

[17] Anxin Wu, Guorui Feng. Payload Mismatch Detection of Image Steganalysis Using Ensemble Linear Discriminant Clustering[J]. IEEE Trans. on Signal Processing, 2015: 1-4.

[18] Yu Dong, Tao Zhang, Xiaodan Hou, et al. A New Steganalysis Paradigm Based on Image Retrieval of Similar Image-Inherent Statistical Properties and Outlier Detection[J]. Wireless Communications & Signal Processing (WCSP), 2015: 1-5.

Improvement of Image Universal Blind Detection based on Training Set Construction

DUAN Huai-feng[1], YANG Yu[2], LEI Min[2], WANG Hui-hua[1]

(1. School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China; 2. Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The practicability of existing universal blind detection reduced greatly due to the generalization problem. According to the principle of orthogonal design, this paper builds three sample sets of embedding rates mismatch, embedding algorithms mismatch and image sources mismatch between the training sample and the testing sample. The three sets are used to test the detection error rates of SPAM and Rich Model in the case of embedding rates mismatch, embedding algorithm mismatch and image source mismatch. This paper proposed several methods to improve the generalization ability of the universal blind detection, including training the sample by small embedding rates, learning various kinds of embeddingalgorithms, pre-classifying the testing sample and improving the IQM algorithm. The results show that the that the performance of the improved algorithm is significantly improved.

Key words: universal blind detection; generalization ability; SPAM; rich model; practicability