

基于多残差马尔科夫模型的图像拼接检测

罗 霄 阳 霍 宏 涛 王 梦 思 陈 亚 飞

(中国人民公安大学信息技术与网络安全学院 北京 100038)

摘 要 针对传统马尔科夫特征计算差值矩阵的方式单一、拼接检测鲁棒性不强的问题,提出彩色多残差马尔科夫特征拼接检测模型。该模型引入隐写检测模型(Rich Models for Steganalysis,SRM)中的多种残差类型来改进传统马尔科夫特征,从 R,G,B 3 个通道分别提取 10 种不同类型的马尔科夫特征,训练 30 个独立的 SVM 分类器,最后通过决策判断进行分类预测。该方法在哥伦比亚大学彩色拼接检测库上达到了 95.40% 的准确率。

关键词 图像取证,拼接检测,隐写检测模型(SRM),马尔科夫特征

中图分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.04.029

Passive Image-splicing Detection Based on Multi-residual Markov Model

LUO Xiao-yang HUO Hong-tao WANG Meng-si CHEN Ya-fei

(Institute of Information Technology and Network Security, People's Public Security University of China, Beijing 100038, China)

Abstract Aiming at the problem that calculating the difference matrix singly for the traditional Markov feature is not robust to the splicing detection, this paper presented a color multi-residual type Markov feature for splicing detection. This method introduces the residual model from rich models for steganalysis(SRM) to improve the traditional Markov features, respectively extracts 10 different types of Markov features from three color channels, and trains 30 unique SVM classifiers to make the classification through the proposed decision-making algorithm. This method achieves the accuracy of 95.40% at Columbia image splicing detection evaluation dataset.

Keywords Image forensics, Splicing detection, Rich models for steganalysis, Markov feature

1 引言

图像篡改在日常生活中随处可见,网友在网上通过图像的拼接、复制、粘贴来恶搞名人,甚至有不法分子通过恶意篡改图像来进行敲诈勒索,影响极其恶劣。数字图像作为电子证据的重要组成部分,在刑事诉讼、民事纠纷中扮演着重要的角色;在判断图像合法性的过程中,不仅要在提取的过程中保证图像的完整性不被破坏,更要保证图像的真实性。对此,许多研究人员提出了不同的图像篡改检测方法来检测图像的真实性,这些检测方法分为两种类型:主动检测和被动检测。主动检测指在图像生成的同时在图像中添加额外的信息来保证图像的真实性,如数字水印、数字签名,一旦图像被篡改,水印和签名就会被破坏。被动检测指在不需要任何先验知识的前提下,通过图像中的统计特征等对图像进行篡改检测。该方式在实际工作中具有更高的应用价值,例如在对互联网上传播的图像进行取证时,只能依靠被动检测的方式判断图像的真伪。在众多图像篡改操作中,图像拼接是篡改程度最大的操作之一,它将一幅图像的一部分复制到另一幅图像的某个区域上,以改变原图像的意义,从而造成误解。图像拼接操作简单,而且经过后期处理(如 jpeg 压缩、中值滤波等)后篡改

的痕迹不易被发现,因此拼接检测在被动取证中占有很重要的地位。Farid 等人^[1]利用小波系数矩特征进行拼接检测。Fu 等人^[2]引入 Hilbert-Huang(H-H)变换特征进行拼接检测。刘晓霞等人^[3]提出了一种基于韦伯局部特征的图像拼接检测,对待测图像进行离散余弦变换后得到 DCT 系数矩阵,进而提取 WLD 特征,以刻画图像的频率分布特性;该算法的不足之处在于准确率不是很高。Shi 等人^[4]首次提出将 markov 特征引入图像拼接检测中,对图像进行块 DCT 变换后计算马尔科夫特征,将转移概率矩阵作为特征向量并使用支持向量机进行训练,取得了较好的检测效果。He 等人^[5]对马尔科夫特征进行了拓展,分别从 DCT 域和 DWT 域提取 markov 特征作为特征向量,并使用主成分分析方法对特征向量进行降维。Su 等人^[6]提出将空间域的游程直方图与频率域马尔科夫特征相结合,经过 PCA 降维后使用关联向量机进行分类检测;但是算法的时间复杂度过高。袁金桥^[7]提出了“高阶子带适合提取拼接信息”的观点,通过拟合系数分布的函数模型计算观测值概率,并根据概率划分将系数映射为马尔科夫状态;该算法的缺点是对不同拼接数据集进行检测的鲁棒性不强。李燕等人^[8]提出了一种基于纹理描述的图像拼接篡改检测方法,即将局部二值模式(LBP)和共生矩阵相结

到稿日期:2017-02-07 返修日期:2017-05-12 本文受公安部技术研究计划项目(2014JSYJB007)资助。

罗霄阳(1992—),男,硕士生,主要研究方向为网络安全、模式识别,E-mail:luoge50@sina.com;霍宏涛(1972—),男,博士,教授,主要研究方向为网络安全、模式识别,E-mail:huotony@163.com(通信作者);王梦思(1991—),女,硕士生,主要研究方向为网络安全、模式识别。

合,在残差图像的基础上利用改进的 LBP 技术对图像纹理进行描述,然后借助两类共生矩阵获取不同的图像特征,再通过 SVM 进行训练分类;其不足之处在于特征组合的效率不高。

2 改进的马尔科夫模型

马尔科夫特征是经典的概率特征,转移概率矩阵反映了图像中各个像素点间的隐含关系。由于在图像拼接的过程中引入了外部元素,造成自身统计特征的改变,从而导致转移概率矩阵发生变化。然而,几乎所有的检测算法在对彩色图像进行检测时都将彩色图像转换成灰度图像,然后再进行检测,这在一定程度上改变了原始图像的统计特征,掩盖了拼接的痕迹,导致检测效果不佳;而且大多数算法的特征较为单一,只提取了像素邻域内特定方向的特征,对不同类型的拼接图像的鲁棒性不强。针对上述不足,本文提出基于彩色多残差马尔科夫特征的图像拼接检测方法。基于图像隐写检测提出的残差模型^[10],针对不同残差模型计算不同方向的转移概率矩阵,并对 R,G,B 3 个通道分别提取马尔科夫特征进行拼接检测,避免了将彩色图像转换成灰度图像带来的问题。

2.1 SRM 模型简介

SRM 是 Jessica 等人^[10]提出的隐写检测模型。该模型由许多子模型的集合构建而成,这种子模型由邻接的联合分布模型组成。通过结合联合分布模型与不同的线性和非线性高通滤波器来计算图像中不同类型的噪声残差,最后计算噪声残差矩阵的灰度共生矩阵,并将其进行量化后作为隐写检测的特征向量。SRM 中有 9 类残差矩阵,分别是:1st,2nd,3rd,EDGE3×3 和 EDGE5×5,其中 1st,3rd,EDGE3×3 和 EDGE5×5 根据残差方向的不同又各分为两种类型。每种残差矩阵与不同的 spam 和 minmax 滤波器相结合,从而得到 39 种模型,如图 1 所示。

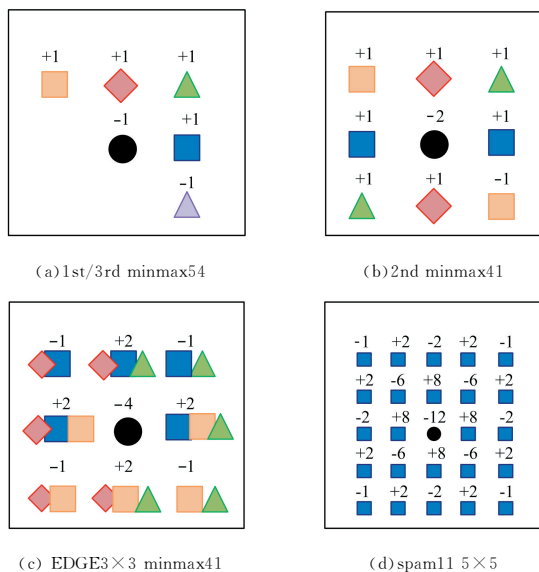


图 1 部分残差类型

Fig.1 Partial residual types

2.2 基于 SRM 残差的改进马尔科夫模型

传统马尔科夫模型分别构造图像的水平、垂直、主对角

线、次对角线 4 个方向的残差矩阵,分别求出 4 个残差矩阵不同方向的转移概率矩阵并将其作为拼接检测的特征向量。残差矩阵的生成只与步长为 1 的相邻像素存在线性关系,对拼接检测的鲁棒性不强。对此,我们引入上述 SRM 模型中的 9 种残差矩阵的计算模型,以扩展马尔科夫特征的多样性。SRM 模型中的 9 种残差模型不仅扩展了残差步长(见图 1(d)),还扩大了残差的计算域。传统马尔科夫模型中残差矩阵的值只与原始矩阵单一方向的邻域值有关,而 SRM 模型中残差矩阵的值受原始矩阵多个方向邻域的影响,如图 1(b)所示,中心点垂直方向上的残差值等于上下邻域值之和减去二倍中心点的值。通过多种残差矩阵的计算,能够更好地将拼接操作对图像内部的扰动凸显出来。分别计算每个颜色通道的 10 种残差矩阵,即 9 种残差类型以及传统马尔科夫残差矩阵模型,最后得到 R,G,B 3 个通道共 30 种残差矩阵,并计算每个残差矩阵的马尔科夫特征(转移概率矩阵)。计算框架如图 2 所示(以蓝色通道为例)。

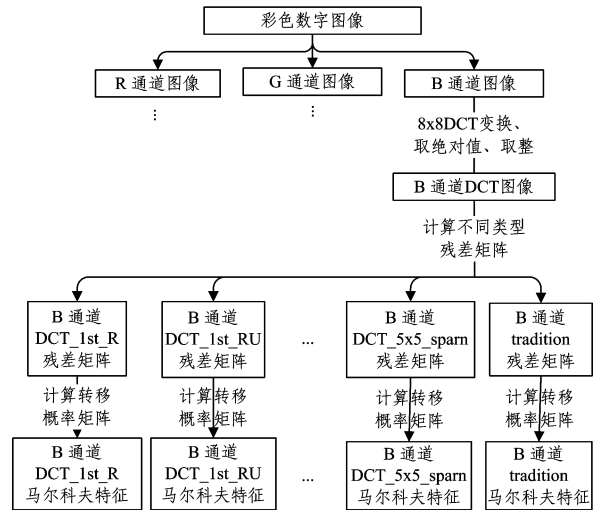


图 2 多残差马尔科夫特征提取框架

Fig.2 Multi-residual Markov feature extraction framework

多残差马尔科夫特征的计算过程分为以下 5 步。

第 1 步 将彩色图像 I 分成 R,G,B 3 个矩阵, I 的大小为 $M \times N \times 3$, R,G,B 的大小皆为 $M \times N$ 。

第 2 步 对每个矩阵分别进行块大小为 8×8 的离散余弦变换,得到 3 个对应的 DCT 系数矩阵 F_r,F_g 和 F_b ,对 3 个矩阵取绝对值,之后再取整。

第 3 步 计算残差矩阵。将 10 个残差矩阵分别命名为 1st_R,1st_RU,2nd_H,3rd_R,3rd_RU,3×3_R,3×3_spam,5×5_R,5×5_spam,tradition。每种残差类型产生 4 个方向(上、下、左、右或水平、垂直、主对角线、次对角线)的残差矩阵,其中 3 个残差矩阵 1st_R,3rd,RU 和 tradition 的计算方法如下所示(以下计算均以蓝色通道矩阵 F_b 为例,其中 $I \in [1,M],J \in [1,N]$,其余详见文献^[10]。

1st_R:

$$Fb_r(I,J) = Fb(I,J+1) - Fb(I,J) \tag{1}$$

$$Fb_l(I,J) = Fb(I,J-1) - Fb(I,J) \tag{2}$$

$$Fb_u(I,J) = Fb(I-1,J) - Fb(I,J) \tag{3}$$

$$Fb_d(I, J) = Fb(I+1, J) - Fb(I, J) \quad (4)$$

3rd_RU:

$$Fb_m(I, J) = -Fb(I-2, J+2) + 3 * Fb(I-1, J+1) - 3 * Fb(I, J) + Fb(I+1, J-1) \quad (5)$$

$$Fb_m(I, J) = -Fb(I-2, J-2) + 3 * Fb(I-1, J-1) - 3 * Fb(I, J) + Fb(I+1, J+1) \quad (6)$$

$$Fb_{nd}(I, J) = -Fb(I+2, J+2) + 3 * Fb(I+1, J+1) - 3 * Fb(I, J) + Fb(I-1, J-1) \quad (7)$$

$$Fb_{nd}(I, J) = -Fb(I+2, J-2) + 3 * Fb(I+1, J-1) - 3 * Fb(I, J) + Fb(I-1, J+1) \quad (8)$$

tradition:

$$Fb_h(I, J) = Fb(I, J) - Fb(I+1, J) \quad (9)$$

$$Fb_v(I, J) = Fb(I, J) - Fb(I, J+1) \quad (10)$$

$$Fb_d(I, J) = Fb(I, J) - Fb(I+1, J+1) \quad (11)$$

$$Fb_m(I, J) = Fb(I+1, J) - Fb(I, J+1) \quad (12)$$

第 4 步 引入阈值 T , 将矩阵中所有元素的范围约束在 $[-T, T]$ 之间。若残差矩阵(如 $Fb_r(I, J)$)中有元素大于 T , 则将其设为 T ; 若小于 $-T$, 则将其设为 $-T$ 。本文将阈值 T 设为 4。

第 5 步 计算每个残差矩阵的转移概率矩阵, 根据不同类型的残差矩阵, 采用两种不同方向的转移概率矩阵计算方式: 1st_R, 3rd_R, 3x3_R 和 5x5_R 采用式(13)–式(16)进行计算, 1st_RU, 2nd_H, 3rd_RU, 3x3_spam, 5x5_spam 和 tradition 采用式(17)–式(20)进行计算。

$$P1_r \{Fb_r(I, J+1) = y | Fb_r(I, J) = x\} = \frac{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x, Fb_r(I, J+1) = y)}{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x)} \quad (13)$$

$$P1_l \{Fb_r(I, J-1) = y | Fb_r(I, J) = x\} = \frac{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x, Fb_r(I, J-1) = y)}{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x)} \quad (14)$$

$$P1_u \{Fb_r(I-1, J) = y | Fb_r(I, J) = x\} = \frac{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x, Fb_r(I-1, J) = y)}{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x)} \quad (15)$$

$$P1_d \{Fb_r(I+1, J) = y | Fb_r(I, J) = x\} = \frac{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x, Fb_r(I+1, J) = y)}{\sum_{I=2}^{M-1} \sum_{J=2}^{N-1} \delta(Fb_r(I, J) = x)} \quad (16)$$

$$P2_h \{Fb_h(I+1, J) = y | Fb_h(I, J) = x\} = \frac{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_h(I, J) = x, Fb_h(I+1, J) = y)}{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_h(I, J) = x)} \quad (17)$$

$$P2_v \{Fb_v(I, J+1) = y | Fb_v(I, J) = x\} = \frac{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_v(I, J) = x, Fb_v(I, J+1) = y)}{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_v(I, J) = x)} \quad (18)$$

$$P2_d \{Fb_d(I+1, J+1) = y | Fb_d(I, J) = x\} = \frac{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_d(I, J) = x, Fb_d(I+1, J+1) = y)}{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_d(I, J) = x)} \quad (19)$$

$$P2_m \{Fb_m(I, J+1) = y | Fb_m(I+1, J) = x\} = \frac{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_m(I+1, J) = x, Fb_m(I, J+1) = y)}{\sum_{I=1}^{M-1} \sum_{J=1}^{N-1} \delta(Fb_m(I, J) = x)} \quad (20)$$

其中, $x, y \in \{-T, -T+1, \dots, 0, \dots, T\}$, $\delta(A=x, B=y) = \begin{cases} 1, & \text{if } A=x \text{ and } B=y \\ 0, & \text{Otherwise} \end{cases}$ 。图 3 给出了当 $T=3$ 时 Fb_r 向上的

转移概率矩阵 $P1_u$, 矩阵中的数值 0.510 表示该矩阵 Fb_r 中值为 -1 的点向上转移成值为 1 的点的概率为 0.510。最终, 所有生成的转移概率矩阵都将作为特征向量进行拼接检测。每个颜色通道的特征向量维数为 $(2T+1) \times (2T+1) \times 4 \times 10$ 。

	-3	-2	-1	0	1	2	3
-3	0.184	0.025	0.013	0.008	0.031	0.130	0.696
-2	0.058	0.037	0.023	0.017	0.086	0.398	0.119
-1	0.110	0.117	0.097	0.094	0.510	0.247	0.073
0	0.243	0.432	0.527	0.578	0.223	0.119	0.041
1	0.164	0.228	0.206	0.182	0.078	0.053	0.025
2	0.092	0.084	0.063	0.052	0.029	0.024	0.014
3	0.149	0.077	0.072	0.068	0.042	0.030	0.031

图 3 $T=3$ 时 Fb_r 向上的转移概率矩阵 $P1_u$

Fig. 3 Fb_r upward transition probability matrix $P1_u$ when $T=3$,

2.3 拼接检测框架

在拼接检测中, 对每一张图像分别提取 3 个颜色通道, 共 30 种马尔科夫特征。在 SVM 训练阶段, 不对这 30 种特征进行融合, 而是利用 PCA 将每一种特征单独进行降维, 当 $T=4$ 时, 每种特征的维数为 $(2 \times 4 + 1) \times (2 \times 4 + 1) \times 4 = 324$, 降维后每种特征的维数为 50。然后, 用独立的 SVM 对每种降维后的特征进行训练, 共训练 30 个 SVM 分类器, 这样可以将每种特征的优势充分发挥出来。在测试阶段, 使用训练好的 SVM 对每一个特征进行分类。拼接检测框架如图 4 所示。

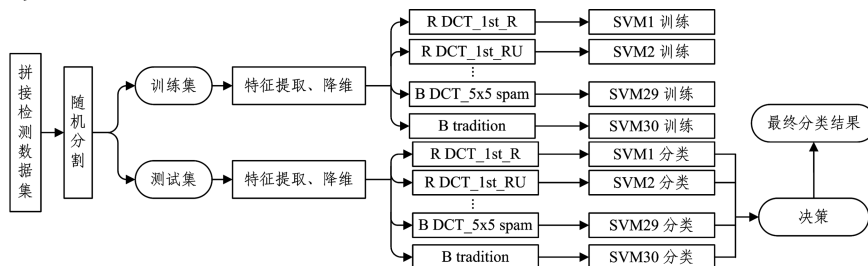


图 4 拼接检测框架

Fig. 4 Splicing detection framework

在决策过程中,把30种特征分为10类,每一类特征有R,G,B 3个通道的子特征。如果每一类特征在R,G,B 3个通道的两个以上中被判断为拼接,则该图像在该特征下被认定为拼接图像。例如,对于某幅图像,R通道的DCT₅×5 spam、G通道的DCT₅×5 spam和B通道的DCT₅×5 spam这3个特征中有两个以上的SVM分类器认定其为拼接图像,则DCT₅×5 spam特征判断该图像为拼接图像。对于某张图像,若有6类以上的特征判断其为拼接图像,则认定它为拼接图像。

3 实验与结果分析

本节将对实验的数据集进行介绍,基于不同的参数对提出的拼接检测模型进行测试,以获取最佳的参数值,并将该检测模型与其他检测算法进行对比。图像篡改者通常会使用多种后期处理手段来掩盖拼接痕迹。在实验中,对拼接图像进行中值滤波、JPEG压缩等操作来模拟图像篡改者的后期处理,以测试算法的鲁棒性。实验环境为Matlab2015a,实验机器为Intel(R) core(TM)i5-3317U CPU1.7GHz,支持向量机为Matlab自带的SVM工具箱;使用RBF核函数,设置RBF_sigma为6.5。

3.1 数据集的选择

传统的拼接检测数据集包括中国科学院的CASIA数据集、哥伦比亚大学的灰度拼接检测数据集和哥伦比亚大学的彩色拼接检测数据集。本文的检测模型基于彩色图像,因此选择哥伦比亚大学的彩色拼接检测数据集^[13]进行测试。该测试集中包含183张真实图像和180张拼接图像,图像大小范围从757×568到1152×768,格式为未压缩的TIFF或BMP。在创建拼接图像时使用了真实的图像并且没有任何后期处理,真实的图像中包含完整的EXIF信息。我们共进行了20次随机测试,每次测试随机选择数据集中5/6的图像作为训练集,剩下1/6的图像作为测试集,取20次测试的平均值作为结果。实验中主要使用TPR,TNR和ACC 3个参数来衡量算法的性能。真正类率(True Positive Rate, TPR)的计算公式为: $TPR = TP / (TP + FN)$,其刻画的是分类器所识别出的拼接图像占所有拼接图像的比例;真负类率(True Negative Rate, TNR)的计算公式为: $TNR = TN / (FP + TN)$;准确率(Accuracy, ACC)的计算公式为: $ACC = (TP + TN) / (TP + FN + FP + TN)$ 。

3.2 阈值 T 与 PCA 降维维数 n 的选择

阈值 T 决定了转移概率矩阵的大小,同时也影响主成分分析的维数选择。如果 T 过小,则无法通过转移概率矩阵捕获DCT系数中的关联性来区分真实图像与拼接图像;如果 T 过大,则计算复杂度会增加。主成分分析(PCA)提取了特征向量的主要成分,减少了特征向量中无关分量的干扰,PCA维数的选择同样影响分类的准确率。因此,通过实验选择出最佳的阈值 T 与维数 n 。实验结果如表1所列,当 $T=4$ 且 $n=50$ 时,准确率达到了最高(95.4%)。

表1 在不同阈值 T 与维数 n 的情况下的实验结果

Table 1 Experimental results with different threshold T and dimension n

阈值 T	维数 n	TPR/%	TNR/%	ACC/%
3	30	99.00	88.79	93.65
3	50	98.33	89.85	93.89
3	80	96.00	87.88	91.75
3	150	83.33	75.15	79.05
4	30	98.67	88.48	93.33
4	50	98.00	93.03	95.40
4	80	97.00	90.61	93.65
4	150	84.00	79.39	81.59
5	30	99.00	90.91	94.76
5	50	99.33	91.52	95.24
5	80	98.00	86.67	92.06
5	150	83.00	78.79	80.79

表2列出了在一次SVM测试中10个特征对测试集图像的判别结果。本次测试最终的分类结果为 $ACC=93.65\%$, $TPR=96.67\%$, $TNR=90.91\%$,与子特征1和子特征7的精度相同,表明该决策算法可以取得子分类器的最高精度。多次实验表明,不同子特征在随机测试集中的表现性能不同,即不同子特征对不同图像进行拼接检测的灵敏度不同。由于特征与特征之间存在不一致性,即在主成分分析中协方差矩阵的特征值的重要性(大小排序)不同,采用传统的特征融合方法并不能有效提高拼接检测的准确率。通过2.3节中的决策层融合可以充分发挥每个特征的特性,从而提升拼接检测的鲁棒性。

表2 10个子特征在测试集上的性能

Table 2 Performance of 10 sub-features on test set

特征名称	TPR	TNR	ACC
1st_R	96.67	90.91	93.65
1st_RU	93.33	90.91	92.06
2nd_H	96.67	84.85	90.48
3rd_R	93.33	87.88	90.48
3rd_RU	93.33	81.82	87.30
3x3_R	90.00	90.91	90.48
3x3_spam	96.67	90.91	93.65
5x5_R	86.67	78.79	82.54
5x5_spam	90.00	81.82	85.71
Tradition Markov	93.33	90.91	92.06

表3与图5给出了多残差马尔科夫特征与小波系数矩特征、DCT域马尔科夫特征和高频小波子带特征的比较结果。可以看出,本文所提方法的 TPR 和 ACC 明显优于前几种算法,但是文献[12]中的方法的 TNR 高达95.40%,优于本文方法。

表3 在哥伦比亚彩色图像库上4种特征识别结果的对比

Table 3 Comparison of four methods' performance on Columbia image splicing detection evaluation dataset

提取的特征	TPR	TNR	ACC
小波系数矩特征 ^[1]	93.60	89.70	91.60
DCT域 markov ^[11]	83.20	86.80	85.00
高频小波子带 ^[12]	92.00	95.40	93.70
本文提取的特征	98.00	93.03	95.40

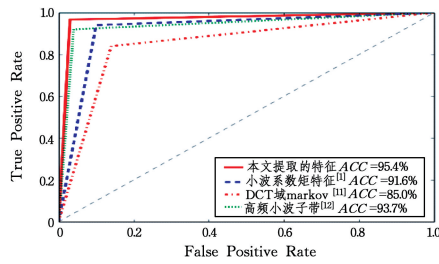


图 5 4 种特征的 ROC 曲线的对比

Fig. 5 Comparison of four features' ROC curve

3.3 鲁棒性测试

为了测试本文拼接检测方法的鲁棒性,将拼接检测数据集中所有的图像进行不同程度的压缩,20 次实验的结果如表 4 所列。可以发现,当压缩质量为 90% 时,准确率达到 100%,而且 10 个子特征的准确率也为 100%,这可能是由于 JPEG 压缩减小了 DCT 变换后高频分量的影响,导致拼接特征被突显。当压缩质量降低至 70% 和 50% 时,仍然取得了 90% 左右的分类准确率,说明本算法对 JPEG 压缩具有很好的鲁棒性。将所有的测试集图像进行中值滤波处理后再进行测试时,准确率略有下降,真正类率有所提升,而真负类率平均下降了 7 个百分点(见表 5),说明经过中值滤波后把真实图像误判为拼接图像的概率有所提升,但是算法的性能基本保持稳定。

表 4 JPEG 压缩图像的拼接检测

Table 4 Splicing detection of JPEG compression image (单位:%)

压缩质量	TPR	TNR	ACC
原始图像	98.00	93.03	95.40
90%	100.00	100.00	100.00
70%	96.33	84.70	90.24
50%	98.00	81.82	89.52

表 5 中值滤波图像的拼接检测

Table 5 Splicing detection of median filter image (单位:%)

滤波块大小	TPR	TNR	ACC
原始图像	98.00	93.03	95.40
3×3	98.67	86.82	92.46
5×5	99.33	86.82	92.78

结束语 本文提出了基于 SRM 残差马尔科夫特征的图像拼接检测算法,该算法通过提取 3 个彩色通道中不同类型残差矩阵中的马尔科夫特征进行拼接检测,在哥伦比亚彩色拼接检测数据库上达到了 95.4% 的准确率;不足之处在于,由于构造的残差矩阵过多,算法在特征提取阶段的时间复杂度较高。在以后的研究中,将进一步改进该拼接检测模型,优化残差矩阵的选取过程,减少特征提取的种类,在提升算法准确性的基础上减少特征提取的时间。由于在进行 JPEG 鲁棒性测试的过程中,当压缩质量为 90% 时,该算法在 20 次随机测试中对所有测试图像均做出了正确的分类判断。因此,在以后的研究中将对 JPEG 压缩中使用不同块大小的 BDCT 变换和不同的量化矩阵对图像拼接检测中提取的特征是否存在影响进行研究,从而取得更高的拼接检测准确率。

参考文献

[1] FARID H,LYU S.Higher-order Wavelet Statistics and their Application to Digital Forensics[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003.

[2] FU D,SHI Y Q,SU W. Detection of image splicing based on Hilbert-Huang transform and moments of characteristic functions with wavelet decomposition[C]// Proceedings of the 5th International Workshop on Digital Watermarking, Berlin: Springer,2006:177-187.

[3] LIU X X,LI F,XIONG B. Image splicing detection using Weber local descriptors[J]. Computer Engineering and Applications, 2013,49(12):140-143. (in Chinese)

刘晓霞,李峰,熊兵. 基于韦伯局部特征的图像拼接检测[J]. 计算机工程与应用,2013,49(12):140-143.

[4] SHI Y Q,CHEN C,CHEN W,et al. A natural image model approach to splicing detection[C]//The Workshop on Multimedia & Security. ACM,2007:51-62.

[5] HE Z,LU W,SUN W,et al. Digital image splicing detection based on Markov features in DCT and DWT domain[J]. Pattern Recognition,2012,45(12):4292-4299.

[6] SU B,YUAN Q,ZHANG Y,et al. Rapid Image Splicing Detection Based on Relevance Vector Machine[M]// Digital Forensics and Watermarking. Springer Berlin Heidelberg, 2012: 300-310.

[7] YUAN Q Q. Image Splicing Blind Detection Based on Enhanced MarkovModel in DWT Domain[D]. Shanghai: Shanghai Jiao Tong University,2014. (in Chinese)

袁全桥. 基于小波域改进马尔科夫模型的图像拼接盲检测研究[D]. 上海:上海交通大学,2014.

[8] LI Y,ZHONG L,LI J. Detection of image splicing forgery based on LBP and co-occurrence matrix[J]. Journal of Wuhan University (Natural Science Edition), 2015, 61(6): 517-524. (in Chinese)

李燕,钟磊,李健. 基于 LBP 和共生矩阵的图像拼接篡改检测[J]. 武汉大学学报(理学版),2015,61(6):517-524.

[9] QIU X,LI H,LUO W,et al. A universal image forensic strategy based on steganalytic model[C]//Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security. ACM,2014:165-170.

[10] FRIDRICH J,KODOVSKY J. Rich Models for Steganalysis of Digital Images[J]. IEEE Transactions on Information Forensics & Security,2012,7(3):868-882.

[11] BATTISTI F,CARLI M,NERI A. Image forgery detection by means of no-reference quality metrics[C]// Media Watermarking,Security,and Forensics. International Society for Optics and Photonics,2012:265-298.

[12] YUAN Q Q,SU B,ZHAO X D,et al. Image splicing detection based on high frequency wavelet Markov features[J]. Journal of Computer Applications,2014,34(5):1477-1481. (in Chinese)

袁全桥,苏波,赵旭东,等. 基于高频小波子带马尔科夫特征的图像拼接检测[J]. 计算机应用,2014,34(5):1477-1481.

[13] Columbia DVMM Research Lab(2004) Columbia image splicing detection evaluation dataset[OL]. <http://www.ee.columbia.edu/ln/dvmm/downloads/authspluncmp>.