

# 合成图像被动取证技术研究进展

王 伟<sup>1</sup>, 曾 凤<sup>1</sup>, 章国安<sup>1\*</sup>, 段新涛<sup>2</sup>

(1. 南通大学 电子信息学院, 江苏 南通 226019; 2. 河南师范大学 计算机与信息技术学院, 河南 新乡 453007)

**摘要:**合成图像取证是数字图像取证的一个重要研究内容, 研究合成图像被动取证技术对确保数字图像的原始性和真实性有着极其重要的意义. 根据合成图像被动取证技术研究角度不同, 将现有的技术方法分为 3 大类: 基于传统特征的合成图像取证、基于盲源分离的合成图像取证、基于稀疏表示的合成图像取证. 然后分别介绍了 3 类取证技术的基本特征和典型算法, 并对不同方法进行了分析比较和总结. 最后指出当前研究中存在的一些主要问题, 给出本领域未来的研究方向.

**关键词:** 图像合成; 被动取证; 篡改检测; 多媒体安全

中图分类号: TN911.73

文献标志码: A

文章编号: 1673-2340(2013)03-0001-06

## Survey on Techniques of Composite Image Passive Forensics

WANG Wei<sup>1</sup>, ZENG Feng<sup>1</sup>, ZHANG Guo-an<sup>1\*</sup>, DUAN Xin-tao<sup>2</sup>

(1. School of Electronics and Information, Nantong University, Nantong 226019, China;

2. School of Computer and Information Technology, Henan Normal University, Xinxiang 453007, China)

**Abstract:** Composite image forensics is a vital research area of digital image forensics. The research on techniques of composite image passive forensics is important to ensure the originality and authenticity of digital image. According to different research areas, the techniques of composite image passive forensics fall into three categories: techniques based on traditional characteristics, techniques based on blind source separation, and techniques based on sparse representation. The basic characteristics, typical methods, as well as comparison and analysis of various algorithms are summarized in detail for each category. The main problems in the current research are pointed out and the future directions are presented.

**Key words:** image composition; passive image forensics; tamper detection; multimedia security

随着图像获取设备(如数码相机)和图像处理软件(如 Photoshop)的飞速发展, 数字化图片在互联网广泛传播, 这也使得图像内容的真实性、完整

性保护工作显得尤为重要. 近年来新闻图片造假案例层出不穷, 给社会造成严重的负面影响. 如华南虎事件、广场鸽事件、藏羚羊事件等, 因此鉴定数字

收稿日期: 2013-06-09

基金项目: 国家自然科学基金项目(61371113, U1204606); 江苏省自然科学基金项目(BK20130393); 江苏省高校自然科学基金项目(12KJB510026, 12KJB510025); 南通大学博士科研启动基金项目

作者简介: 王伟(1983—), 男, 讲师, 博士.

\* 通信联系人: 章国安(1965—), 男, 教授, 博士, 主要研究方向为无线通信和信息安全. E-mail: gzhang@ntu.edu.cn

图像的真伪具有迫切的现实意义。

合成拼接图像检测是目前图像取证研究的重点<sup>[1]</sup>。因为对于技术娴熟的篡改伪造者,仅仅利用合成拼接操作,就可以十分轻松和容易地制作出一幅足够以假乱真的图像,因此,如何检测合成图像的原始性已经成为当前数字图像取证的一个重要研究内容。

### 1 合成图像被动取证技术研究现状

目前针对图像拼接合成的鉴定方法主要有两类:主动取证和被动取证。被动取证仅根据待鉴定图像本身判断其真伪,目前正受到国内外学者的极大关注。现阶段的合成图像被动取证研究主要集中在寻找能够描述不同拼接痕迹的取证特征,但目前基于盲源分离和稀疏表示理论的合成图像被动取证新方法也受到研究者的广泛关注,具体分类如图1所示。

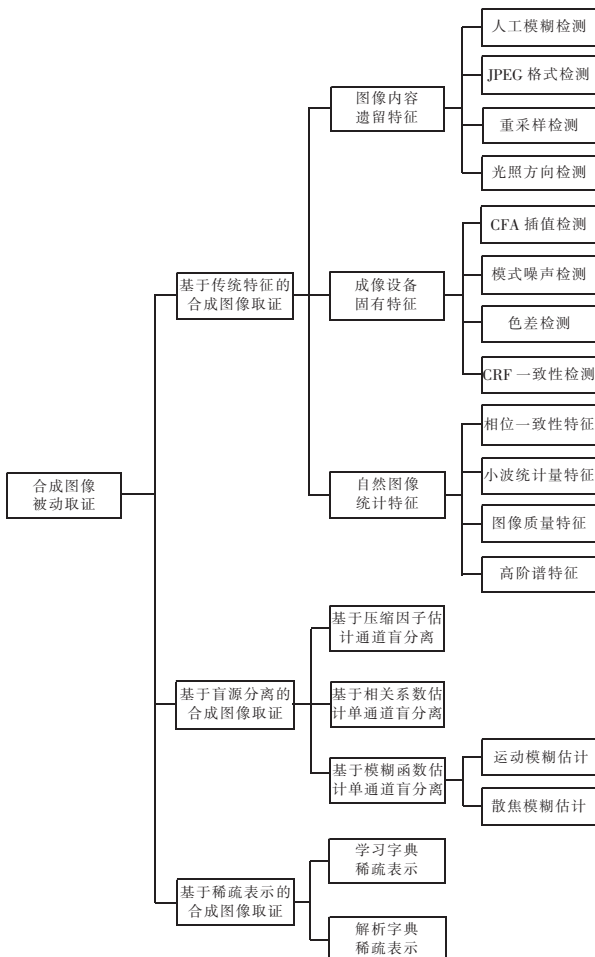


图1 合成图像被动取证方法分类图

#### 1.1 基于传统特征的合成图像取证

该类技术试图选择能够描述拼接痕迹的取证特征来检测图像是否被伪造。检验图像内容中的遗留特征是被取证的重要手段。篡改者为了掩饰生硬的拼接边缘往往会进行一定程度的模糊操作,根据是否存在人工模糊的局部区域从而判断图像中来源不明的拼接内容。Hsiao和Pei<sup>[2]</sup>,王波等<sup>[3]</sup>分别提出基于频域特性和异常色调率的局部模糊检测算法,根据定位出的局部模糊区域来识别篡改痕迹。Sutcu等<sup>[4]</sup>和周琳娜等<sup>[5]</sup>分别提出利用小波变换和同态滤波方法来对图像边缘的模糊程度进行估计,根据估计结果的不一致性来辨识伪造区域。而王睿等<sup>[6]</sup>对上述方法进行了改进,通过计算每个像素所在区域的局部相关性,以此判断该像素是属于模糊区域还是非模糊区域来实现篡改检测与定位。但如果真实图像本身存在模糊图像块,上述的被动取证方法就会产生虚警。当利用两幅或多幅JPEG图像合成一幅新的JPEG图像时,要在内容拼接后进行JPEG编码,也就是说某些内容经过了两次DCT系数量化,根据背景区域和合成区域具有不一致的JPEG量化系数可实现篡改区域的检测。He等<sup>[7]</sup>人通过分析图像DCT系数的双重量化效应,首次实现了JPEG图像的篡改检测和篡改区域的自动定位。Ye等<sup>[8]</sup>,李晟和张新鹏<sup>[9]</sup>则用不同的压缩因子对待检测图像进行尝试性的二次压缩,根据二次压缩在图像局部引入的失真不一致性实现JPEG合成图像的篡改检测。Farid<sup>[10]</sup>提出了一种称为JPEG ghost的现象,利用此现象可以识别不同质量因子的图像进行的拼接。郑二功和平西建<sup>[11]</sup>根据压缩前后篡改区域与非篡改区域失真的不一致性,提出了一种基于图像块效应指数映射图的篡改检测新方法。在由几幅图像通过拼接各自部分形成一幅新的合成图像过程中,为了消除拼接后视觉畸变,通常要对篡改的局部进行一些重采样或插值处理,因此可根据像素之间特殊的统计相关性来察觉这种行为。Popescu和Farid<sup>[12]</sup>,Kirchner和Bohme<sup>[13]</sup>提出利用EM算法来检测由于插值操作所引入的周期特性,通过检测这种周期特性可判断重采样操作的存在。王伟和方勇<sup>[14]</sup>针对上述EM

算法过度依赖参数初值选取的问题,提出了一种利用有限差分算法来检测周期特性的方法. Mahdian 和 Saic<sup>[15]</sup>提出利用方差的周期性来检测图像所经历的插值操作. 为了准确的估计缩放和旋转因子, Wei 等<sup>[16]</sup>人根据像素行向量(或列向量)二阶差分后离散谱中尖峰的归一化频率与缩放倍率的关系,估计出子块的缩放与旋转参数,从而揭示了篡改区域的存在. 在实际应用中,光照条件尤其适合于检测拼接-合成类型的篡改图像,因为不同光照条件图像合成时往往很难使其光照效果保持一致. Johnson 和 Farid<sup>[17]</sup>提出利用图像中各物体的反射分量来估计入射光方向,然后根据入射光方向的吻合情况来识别篡改. 他们还提出利用人眼睛对光线的高亮反光来估计三维光源方向<sup>[18]</sup>. 即使在复杂光照环境下,也可依据图像中的不同对象分别估计光照特性<sup>[19]</sup>. 鉴别这些特性是否一致可用于辨识合成图像,但这些方法大都需要预先知道取证特征,而这样的假设在数字图像被动盲取证的背景下有时候无法满足,因此限制了这类方法的应用环境和条件.

另一方面,检验由图像生成设备引起的固有特征也是识别伪造图像的重要证据. 成像设备会在生成的图像中引入一些设备的固有特征,通过检测设备固有特征的一致性可判断是否发生伪造. 大多数数码相机拍摄的彩色照片是由单个传感器结合色彩滤波阵列(Color Filter Array, CFA)插值得到的,由于 CFA 插值会在色彩通道和原始像素间引入特殊的统计相关性,篡改操作会破坏这种相关性,因此可被察觉. Popescu 和 Farid<sup>[20]</sup>提出使用 EM 算法来对 CFA 插值引入的周期相关性进行检测,根据这种周期相关性是否被破坏来识别篡改. Long 和 Huang<sup>[21]</sup>用一个均方差模型来表示 CFA 插值的周期关联性,然后使用神经网络方法来对真伪图像进行分类. 曹刚等<sup>[22]</sup>人在上述方法的基础上提出了一种新的 CFA 插值表示模型,根据阶跃型边缘内插特征的一致性来辨识图像拼接. 但随着可感应多种色彩的 CMOS 传感器技术的发展,一些新型相机拍摄的图像将不具备这种邻域插值相关性,从而使这类取证方法失效. 数码相机在成像环节中会不可避免

地引入模式噪声,可根据该特定模式噪声的一致性来作为图像拼接的检测依据. Lukas 等<sup>[23]</sup>人提取一定数量样本图像的噪声残差和待检测图像的噪声残差进行比对,采用相似度度量方法对待检测图像进行取证. 针对传感器模式噪声提取过程中易受到图像内容影响的特点, Li 等<sup>[24]</sup>人提出改进的模式噪声提取方法. 同样, Kang 等<sup>[25]</sup>也提出了模式噪声提取的改进方法,他们采用噪声残差频谱的相位谱作为模式噪声,对 JPEG 压缩和纹理复杂图像的认可准确率有较大幅度的提升. 但是,该类方法在使用时需要已知一组该相机拍摄的图像,通过对图像进行训练得到该相机的模式噪声特征. 相机镜头的折射效应会在图像的不同位置产生有规律的轻微色彩失真,这种色差在一幅自然图像内应该是一致的. Johnson 和 Farid<sup>[26]</sup>根据横向色差的一致性来判断图像是否发生篡改. 数码相机的光电转换器件还会对真实光强产生非线性响应,相机不同,其相机响应函数(Camera Response Function, CRF)也不同,可通过检测图像中 CRF 的一致性来判别其篡改历史. Lin 等<sup>[27]</sup>人提出了一种基于相机响应函数的常态性和一致性的篡改检测方法. Hsu 和 Chang<sup>[28]</sup>提出一种基于几何不变量和相机一致性特征来识别真实图像和拼接图像的半自动分类方法. 此外,他们还在原有方法的基础上提出了一种完全自动的拼接图像检测方法<sup>[29]</sup>. 但该类方法都对训练样本的依赖性较强,如果没有对比训练样本数据库的支持,检测算法将无法实施.

此外,自然图像的一致性特征也可用于合成拼接图像检测. 该类方法是一种基于机器学习的方法,通过提取自然图像的一致性特征,建立图像的统计模型,根据训练分类器完成自然图像和篡改图像的检测. 从目前的研究情况来看,现有的方法基本上都是基于图像的高阶统计特性,提取的统计特征包括相位一致性<sup>[30]</sup>、高阶小波统计量<sup>[31]</sup>、图像质量度量<sup>[32]</sup>、马尔可夫转移概率矩阵和 Hilbert-Huang 变换各阶矩<sup>[33]</sup>和双谱特征<sup>[34]</sup>等. 但是,这些方法大都是利用图像的统计特征模型对正常图像和合成图像进行分类,因此只能检测出合成伪造图像,而不能对伪造区域进行定位.

## 1.2 基于盲源分离的合成图像取证

最近几年,国内有部分学者另辟蹊径,针对合成图像取证问题具有的显著特点:原始图像未知(信号源未知)、篡改手段未知(信号混叠方式未知)、只有一幅待检测图像(观测信号只有一个),通过借鉴单通道盲源分离理论对合成图像被动取证展开了初步研究,并取得了一些成果.方勇等<sup>[35]</sup>针对观测信号相互置换的情况,提出将其建模为一个时变的单通道瞬时混合模型,利用特征域的可分性实现置换混叠区域的检测和分离.他们在此时变混合模型的基础上,针对二维合成图像信号的固有特性,提出将合成图像的篡改区域和未篡改区域等效为不同源图像,建立了合成篡改图像的单通道混合模型.并在此模型基础上提出了一系列基于参数估计的合成图像单通道盲分离方法<sup>[36-39]</sup>,通过估计压缩因子、相关系数和模糊函数等参数来估计混合矩阵,从而实现篡改区域和未篡改区域的分离.

## 1.3 基于稀疏表示的合成图像取证

针对上述建模的单通道合成图像数学模型与传统的叠加混合盲分离模型在混合方式上完全不同——合成图像在任何时刻只有一个源图像分量处于激活状态,段新涛和方勇<sup>[40-41]</sup>提出能否针对合成拼接图像,通过设计合适的字典,通过字典得到其稀疏表示,使得篡改区域图像与未篡改区域图像在稀疏域上具有差异性,利用这种差异性来检测和分离伪造图像,而不是直接估计参数.文献[40]针对一类含有不同背景噪声的合成图像,提出了一种基于K-SVD字典学习的合成图像盲检测方法.该方法首先通过K-SVD算法对合成图像进行训练得到其稀疏表示字典,然后利用学习得到的字典对背景噪声进行去除,最后根据去噪前后图像相减实现篡改区域的检测与定位.文献[41]针对一类由光滑图像和纹理图像拼接而成的合成图像,分别选取Contourlet变换基字典和Local DCT基字典作为稀疏字典.由于合成图像中的光滑区域和纹理区域仅在各自的字典上能够稀疏表示,而在非相应的字典上不能稀疏表示,因此可以把篡改区域和未篡改区域的分离问题转化为求解合成图像的最稀疏表示,根据求得的最稀疏解重构出篡改区域和未篡改区域.

## 2 分析与讨论

综合起来,虽然目前基于传统特征的综合图像取证研究正在蓬勃开展和取得阶段性成果,但仍存在一些亟待发展之处:1)已有的取证方法往往在某些具体条件下才可奏效,而伪造图像情况各异,未必满足特定条件;2)已有的取证方法在察觉单一伪造行为时比较有效,而对鉴别多种篡改联合伪造的图像往往无能为力;3)大多数研究致力于发展用于取证的具体特征和具体的取证方法,而理论研究和模型建立尚显薄弱.

从盲源分离角度研究异源图像合成篡改问题具有很好的新意,特别能为完善其理论体系提供理论基础.但研究者并未能提出比传统取证方法新颖的盲分离算法,由于图像类型的多样性和复杂性,以及图像在获取、传输和处理过程中都可能感染噪声等因素,仅仅依靠对图像区域所受某种操作参数的准确估计来实现合成图像的盲分离往往是不易的.

由于构成合成图像的源图像具有多样性和复杂性,且拼接后的图像往往不会造成视觉上的畸变,现有的稀疏字典构造方法均不能解决这类单通道合成图像的分离问题.而且,原图像只有部分区域被其他图像替换,仍保持很好的内容完整性,可识别性会大大降低.同时,只有一幅待检测图像,这都给设计具有可分性的稀疏字典带来困难.如何根据合成图像中不同来源图像的自身特点以及仅有的先验信息构造有效的稀疏字典,这是篡改区域和未篡改区域能否准确分离中最困难又必须要解决的关键问题.但同时,由于字典在设计时需要考虑几何不变性、局部性、过完备性、方向性等一些因素,为鉴别多种篡改联合伪造的合成图像(如篡改区域经过几何变换、局部模糊和加噪、运动等)提供了可能.

## 3 结束语

合成图像取证是数字图像取证的一个重要研究内容,是多媒体信息安全领域的前沿研究课题.无需任何附加信息,仅根据待鉴定图像本身判断其真伪是一个具有挑战性的难题.现有的合成图像被

动取证可以分为基于传统特征的被动取证、基于盲源分离的被动取证、基于稀疏表示的被动取证等3大类,其中围绕第一类的研究较多。

当前图像被动取证研究仍处于初期。不少检测方法尚待进一步发展以应付更复杂的情况。在我们的研究中注意到,由于构成合成图像的源图像具有多样性和复杂性,我们需要突破现有基于某种特征或条件的“狭义”取证方法,从全新的观点去研究这种单通道合成图像的盲分离机理和方法,建立适用于不同来源图像自身特点的合成图像被动取证理论。

#### 参考文献:

- [1] Farid H. Creating and detecting doctored and virtual images: implications to the child pornography prevention act [R/OL]. (2004-09-30)[2013-05-19]http://www.cs.dartmouth.edu/cms\_file/SYS\_techReport/381/TR2004-518.pdf.
- [2] Hsiao D Y, Pei S C. Detection digital tampering by blur estimation[C]//Proceeding of the First International Workshop on Systematic Approaches to Digital Forensic Engineering, November 7-9, 2005, Taipei, Taiwan. New York:IEEE Xplore, 2005:264-278.
- [3] 王波,孙璐璐,孔祥维,等.图像伪造中模糊操作的异常色调率取证技术[J].电子学报,2006,34(12A):2451-2454.
- [4] Sutcu Y, Coskun B, Sencar H T, et al. Tamper detection based on regularity of wavelet transform coefficients[C]//Proceedings of IEEE International Conference on Image Processing, September 16-Oct 19, 2007, San Antonio, TX. New York:IEEE Xplore, 2007:397-400.
- [5] 周琳娜,王东明,郭云彪,等.基于数字图像边缘特性的形态学滤波取证技术[J].电子学报,2008,36(6):1047-1051.
- [6] 王睿,方勇.一种图像局部模糊检测方法及其在被动图像认证中的应用[J].高技术通讯,2009,19(7):718-723.
- [7] He Junfeng, Lin Zhouchen, Wang Lifang, et al. Detecting doctored JPEG images via DCT coefficient analysis[C]//Proceedings of 9th European Conference on Computer Vision, May 7-13, 2006, Graz, Austria. Berlin:Springer-Verlag, 2006:423-435.
- [8] Ye Shuiming, Sun Q B, Chang E C. Detecting digital image forgeries by measuring inconsistencies of blocking artifact [C]//Proceedings of IEEE International Conference on Multimedia and Expo, July 2-5, 2007, Beijing, China. New York:IEEE Xplore, 2007:12-15.
- [9] 李晟,张新鹏.利用JPEG压缩特性的合成图像检测[J].应用科学学报,2008,26(3):281-287.
- [10] Farid H. Exposing digital forgeries from JPEG ghosts[J]. IEEE Transactions on Information Forensics and Security, 2009, 4(1):154-160.
- [11] 郑二功,平西建.针对一类JPEG图像伪造的被动盲取证[J].电子与信息学报,2010,32(2):394-399.
- [12] Popescu A C, Farid H. Exposing digital forgeries by detecting traces of resampling[J]. IEEE Transactions on Signal Processing, 2005, 53(2):758-767.
- [13] Kirchner M, Bohme R. Hiding traces of resampling in digital image [J]. IEEE Transaction on Information Forensics and Security, 2008, 3(4):582-592.
- [14] 王伟,方勇.基于有限差分的置换图像盲检测方法[J].电子学报,2010,38(10):2268-2272.
- [15] Mahdian B, Saic S. Blind authentication using periodic properties of interpolation[J]. IEEE Transactions on Information Forensics and Security, 2008, 3(3):529-538.
- [16] Wei Weimin, Wang Shuozhong, Zhang Xinpeng, et al. Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(3):507-517.
- [17] Johnson M K, Farid H. Exposing digital forgeries by detecting inconsistencies in lighting[C]//Proceedings of the 7th Workshop on Multimedia and Security, August 1-2, 2005, New York, USA. New York:ACM, 2005:1-10.
- [18] Johnson M K, Farid H. Exposing digital forgeries through specular highlights on the eye[C]//Proceedings of 9th International Workshop on Information Hiding, June 11-13, 2007, Saint Malo, France. Berlin:Springer-Verlag, 2007:311-325.
- [19] Johnson M K, Farid H. Exposing digital forgeries in complex lighting environments[J]. IEEE Transactions on Information Forensics and Security, 2007, 2(3):450-461.
- [20] Popescu A C, Farid H. Exposing digital forgeries in color filter array interpolated images[J]. IEEE Transactions on Signal Processing, 2005, 53(10):3948-3959.
- [21] Long Yangjing, Huang Yizhen. Image based source camera identification using demosaicking[C]// Proceedings of IEEE

- 8th International Workshop on Multimedia Signal Processing, October 3–6, 2006, Victoria, BC. 2006:419–424.
- [22] 曹刚, 赵耀, 倪蓉蓉. 基于边缘 CFA 内插特征一致性的图像拼接检测[J]. 东南大学学报, 2009, 39(3):459–463.
- [23] Lukáš J, Fridrich J, Goljan M. Detecting digital image forgeries using sensor pattern noise[C]//Proceedings of the SPIE, Security, Steganography and Watermarking of Multimedia Contents VIII, January 15–16, 2006, San Jose, California, USA. Bellingham, WA:Society of Photo–Optical Instrumentation Engineers, 2006:362–372.
- [24] Li C T. Source camera identification using enhanced sensor pattern noise[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(2):280–287.
- [25] Kang Xiangui, Li Yinxiang, Qu Zhenhua, et al. Enhancing source camera identification performance with a camera reference phase sensor pattern noise[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2):393–402.
- [26] Johnson M K, Farid H. Exposing digital forgeries through chromatic aberration[C]//Proceedings of the 8th Workshop on Multimedia and Security, September 26–27, 2006, Geneva, Switzerland. New York:ACM, 2006:48–55.
- [27] Lin Zhouchen, Wang Rongrong, Tang Xiaou, et al. Detecting doctored images using camera response normality and consistency[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 20–25, 2005, San Diego, USA. New York:IEEE Xplore, 2005:1087–1092.
- [28] Hsu Y F, Chang S F. Detecting image splicing using geometry invariants and camera characteristics consistency [C]//Proceedings of IEEE International Conference on Multimedia and Expo, July 9–12, 2006, Toronto, Canada. New York:IEEE Xplore, 2006:549–552.
- [29] Hsu Y F, Chang S F. Camera response functions for image forensics:an automatic algorithm for splicing detection[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(4):816–825.
- [30] Chen Wen, Shi Yun Q, Su Wei. Image splicing detection using 2–D phase congruency and statistical moments of characteristic function[C]//Proceedings of SPIE, Security, Steganography and Watermarking of Multimedia Contents IX, January 29–February 1, 2007, San Jose, California, USA. Bellingham, WA:Society of Photo–Optical Instrumentation Engineers, 2007. doi:10.1117/12.704321.
- [31] Sun Shaojie, WU Qiong, LI Guohui. Detection of image compositing based on a statistical model for natural images [J]. Acta Automatica Sinica, 2009, 35(12):1564–1567.
- [32] 张震, 康吉全, 平西建, 等. 用统计特征量实现的图像拼接盲检测[J]. 计算机应用, 2008, 28(12):3108–3111.
- [33] 于振华, 金波, 陈璐艺, 等. 基于马尔科夫和 HHT 的图像拼接盲检测[J]. 信息安全与保密通信, 2010(1):107–108.
- [34] Ananya U, Mudanagudi U. Detection of doctored images using bispectral analysis[C]//Proceedings of International Conference on Image Information Processing, November 3–5, 2011, Himachal Pradesh, India. New York:IEEE Xplore, 2011:1–6.
- [35] 方勇, 王伟, 王睿. 置换混叠信号的盲检测和分离[J]. 应用科学学报, 2009, 27(5):491–497.
- [36] Wang Wei, Fang Yong. Single–channel blind separation of JPEG permuted image using double compression[C]//Proceedings of International Conference on Audio Language and Image Processing, November, 23–25, 2010, Shanghai, China. New York:IEEE Xplore, 2010:439–443.
- [37] 王伟, 方勇. 基于二次模糊相关性的单通道置换图像盲分离[J]. 应用科学学报, 2011, 29(2):169–175.
- [38] Fang Yong, Wang Wei. Blind–restoration–based blind separation method for permuted motion blurred images[J]. Journal of Shanghai University:English Edition, 2011, 15(2):79–84.
- [39] 王伟, 方勇. 单通道散焦模糊置换图像的盲分离[J]. 中国图象图形学报, 2012, 17(1):62–67.
- [40] 段新涛, 方勇. 基于噪声检测的置换混叠图像盲分离[J]. 信号处理, 2011, 27(7):1069–1075.
- [41] 段新涛, 方勇. 基于稀疏分解的置换混叠图像盲分离[J]. 高技术通讯, 2012, 22(4):368–373.

(责任编辑:仇慧)