

Copy-Move Forgeries Detection Based on SIFT Algorithm

Junbin Wang¹, Zhenghong Yang¹, Shaozhang Niu²

¹School of Science, China Agricultural University, Beijing

²School of Computer, Beijing University of Posts and Telecommunications, Beijing

Email: wangjunbin34@163.com

Received: Jul. 28th, 2015; accepted: Aug. 14th, 2015; published: Aug. 19th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In view of the higher time complexity of traditional copy-move forgery detection, bad robustness for the image rotation zoom and other follow-up retouching operations, a copy-move forgeries detection algorithm based on SIFT key points is studied. First, extract SIFT key points on the image by using SIFT algorithm, and then match the key point of extraction. Connect the key points of the match to locate the tampered area. Experimental results show that the algorithm can effectively detect the subsequent tampering of the geometric transformation such as rotation and zoom.

Keywords

Image Forensics, Copy-Move Detection, Scale-Invariant Feature Transform (SIFT), Vector Matching

一种基于关键点的复制粘贴盲检测算法

王珺斌¹, 杨正洪¹, 牛少彰²

¹中国农业大学理学院, 北京

²北京邮电大学计算机学院, 北京

Email: wangjunbin34@163.com

收稿日期: 2015年7月28日; 录用日期: 2015年8月14日; 发布日期: 2015年8月19日

摘要

鉴于传统的复制粘贴篡改检测时间复杂度较高,而且对于图像的旋转缩放等后续润饰操作的鲁棒性不好,误报率高。提出了一种基于SIFT关键点的复制粘贴检测算法。首先利用SIFT算法对整幅图像提取SIFT关键点,再对提取到的关键点进行匹配。将匹配的关键点连接起来,来定位篡改区域。实验效果表明,相比较于传统的复制粘贴检测算法,该算法可以有效地检测出经过旋转缩放等几何变换的后续篡改。

关键词

图像取证, 复制粘贴篡改, SIFT算法, 向量匹配

1. 前言

随着数码相机以及数字扫描设备的普及,数码照片越来越多的出现在人们的生活中,成为了人们日常生活的一部分,并给人们的通讯和交流带来了前所未有的便利。然而由于越来越多的图像处理和编辑软件(如 Photoshop 等)的普及,使得数字图像的篡改变得越来越容易,而篡改的图像如果用于新闻、政治等敏感领域,将会产生极大的社会影响。因此图像取证技术应运而生。对数字图像取证技术的研究发展至今,已有几十年的历史。图像篡改检测的目的就是检测图像的真实性和来源。到目前为止,国内外研究数字图像篡改认证的方法主要分为两大类[1]:主动图像取证技术和被动图像取证技术即图像盲取证技术。图像的主动取证技术主要是基于数字水印和数字签名。数字水印是一种解决数字作品版权问题的技术有许多经典的算法。比如 Scyndel 算法、NEC 算法、及基于 JPEG 和 MPEG 标准的压缩域数字水印算法等等。数字签名技术是指用户用自己的密钥对原始信息进行加密从而形成签名。Schneider 和 Chang 提出了一种基于图像分块的灰度直方图数字签名方法, Luo 和 Liu 提出了一种基于块的灰度均值和量化编码的方法。但是主动取证技术本身存在很大的局限性,例如事先对图像加密或进行数字签名会影响图像的质量、成本较高、存储量较大等。因此不需要对图像进行预处理的被动图像取证技术应运而生。

基于数字图像固有特征的被动盲检测方法是直接根据图像本身进行鉴别的。目前被动盲取证技术仍处在起步阶段,各项技术也处在发展阶段。而其中同幅图像的复制粘贴篡改是最常见的篡改方式之一,也是应用最广泛的图像篡改手段。近些年来针对图像复制粘贴篡改已经出现了相当多的检测方法。早期的图像取证技术的研究热点是针对复制-粘贴的检测技术。考虑到同幅图像的复制-粘贴操作会使得图像中存在两块或者两块以上几乎完全相同的图像区域,所以针对同幅图像的复制-粘贴篡改的取证技术有遍历搜寻法、图像块自相关矩阵法和图像块匹配法。然而,这些算法在检测数据量较大的图像时,算法的时间复杂度和空间复杂度较大,因此,在实际的操作过程中受到了限制。美国的 Fridrich 教授在文献[2]提出了改进方案——从 DCT 量化系数出发寻找相似区域。在 2004 年,美国的 Popescu 教授将主成分分析算法运用到取证中,克服了在检测中遇到的运算量大和图像中存在相似块的难题,并在文献[3]阐述了其研究成果。在国内,来自国防科技大学的李国辉教授等人提出了对图像做 DWT 变换处理后再提取低频图像特征,并用奇异值分解算法(Singular value decomposition, 简称 SVD 算法)对所提取的特征做降维处理,然后再做字典排序进行检测的取证算法[4]。但是已有的算法大都对于传统的“复制-移动-粘贴”模型[5]即复制区域和粘贴区域之间仅仅经过位移变换具有较好的效果,但是由于所提取的特征参数本身的局限性,对于“复制-变换-移动-粘贴”模型[6],这些方法仍然无法有效地检测经过旋转、缩放的复制区域。因此,本文主要提出了一种基于 SIFT 特征点的复制粘贴盲取证算法。即通过提取然后匹配图

像的特征点来对图像的复制粘贴篡改进行检测。由于 SIFT 特征点对于旋转缩放等保持不变性，因此相对于传统的算法，基于特征点的方法对图像经过后续的旋转缩放修饰之后的篡改图像具有更好的效果。

2. 基于关键点的复制粘贴检测算法

2.1. SIFT 关键点简介

一幅图像或者物体的矢量描述，即特征，这对于计算区域复制检测是非常实用的。图像中的特征可以分为全局的和局部的。全局特征通常将一个对象描述为一个整体，局部特征是从对象的特定部分提取出来的。局部特征的提取通常由两个步骤组成。首先，在检测复制区域时，提取的特征对几何变换具有鲁棒性；其次，相同或相近区域的特征具有相似性。

尺度不变特征转换(Scale-invariant feature transform 或 SIFT) [7]-[12]是一种电脑视觉的算法用来侦测与描述影像中的局部性特征，它在空间尺度中寻找极值点，并提取出其位置、尺度、旋转不变量。其应用范围包含物体辨识、机器人地图感知与导航、影像缝合、3D 模型建立、手势辨识、影像追踪和动作比。对已经在被检验过对图像尺度具有不变性，同样对三维视点的变换，压缩，添加噪声，照明的变化方面具有鲁棒性。

SIFT 算法的实质是在不同的尺度空间上查找关键点(特征点)，并计算出关键点的方向。SIFT 所查找到的关键点是一些十分突出，不会因光照，仿射变换和噪音等因素而变化的点，如角点、边缘点、暗区的亮点及亮区的暗点等。

SIFT 算法分为如下四步：

- 1) 尺度空间极值检测：搜索所有尺度上的图像位置。通过高斯微分函数来识别潜在的对于尺度和旋转不变的兴趣点。
- 2) 关键点定位：在每个候选的位置上，通过一个拟合精细的模型来确定位置和尺度。关键点的选择依据于它们的稳定程度。
- 3) 方向确定：基于图像局部的梯度方向，分配给每个关键点位置一个或多个方向。所有后面的对图像数据的操作都相对于关键点的方向、尺度和位置进行变换，从而提供对于这些变换的不变性。
- 4) 关键点描述：在每个关键点周围的邻域内，在选定的尺度上测量图像局部的梯度。这些梯度被变换成一种表示，这种表示允许比较大的局部形状的变形和光照变化。

2.2. 利用 SIFT 算法提取关键点

2.2.1. 在高斯尺度空间寻找极值点

检测尺度空间极值点是通过使用尺度空间(Scale Space)尺度的连续函数在所有可能的尺度内寻找稳定特征以检测出对图像尺度变化具有不变性的位置。高斯卷积核已经被证明是唯一一个可以实现尺度变换的变换核。从高斯差分尺度空间中提取局部极大值或者极小值点以提取关键点：

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

其中 k 是尺度因子，尺度空间图像为 $L(x, y, \sigma)$ ， $*$ 代表卷积运算，高斯 $G(x, y, \sigma)$ 。

二维尺度空间的定义为：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

其中， $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ 表示可变尺度高斯函数， $*$ 是卷积运算， (x, y) 表示空间坐标， σ 表示尺度因子，尺度因子越大，表明平滑程度越好，图像越模糊；尺度因子越小，表明平滑程度越差，图像

越清晰。差分高斯金字塔图像的构建如图 1 所示。

2.2.2. 极值点的精确定位

遍历所有的尺度，记录下图像所处的位置，使用高斯差分公式得到所有可能的特征点，如图 2 所示。在检测尺度空间极值的过程中，将一个像素与其临近的 26 个像素点相互比较，最后精确确定关键点的位置和尺度。

2.2.3. 关键点方向的分配

为了使得到的算子具有旋转不变性，需要在关键点邻域窗口内，统计窗口内的所有像素的梯度方向，选择统计量最大的方向作为关键点的主方向。每个关键点包含三种信息，分别为位置、所处尺度、方向。邻域窗口内每个像素的梯度模值使用式(2)计算，每个像素的方向使用式(3)计算。

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (3)$$

2.2.4. 生成特征点描述子

首先将坐标轴旋转到关键点的方向，以确保旋转不变性。在得到主方向后，将关键点描述成特征向量使关键点数值化。以关键点为中心，取一个大小为 $N \times N$ 大小的窗口，以关键点为中心，将邻域窗口分为 S 块，每块为一个种子点，然后统计每块中的所有像素点在不同梯度方向的模值并累加。图 3 为特征点描述子的构造。

最后得到的 SIFT 关键点 $f(x, y, \sigma, \theta)$ 代表在高斯尺度空间的关键点， (x, y) 代表在图像中的坐标， σ 代表高斯尺度空间的尺度因子， θ 代表主方向。每一个 SIFT 关键点是一个 128 维的特征向量。

2.3. 关键点匹配

关键点匹配即找出具有相同或相似特征的关键点，通常是利用特征向量的欧式距离 d 来衡量特征的相似性。图像复制粘贴检测中的特征匹配通常是利用特征的最近邻 d_1 和次近邻 d_2 之间的比值小于预定的阈值来作为特征相似性度量的准则，即如果满足 $\frac{d_1}{d_2} < T$ 则判断匹配成功。

2.4. 算法框架

- 1) 利用 SIFT 算法提取图像中的 SIFT 关键点，并生成相应的描述向量。
- 2) 把所有描述向量组成一个特征矩阵。
- 3) 找出特征矩阵每一行距离当前关键点欧氏距离最近的关键点及次近的关键点。
- 4) 把最近的关键点的距离除以次近的关键点的距离所得数值与预设阈值进行比较，如果小于阈值则确定这两个点为一对匹配点。
- 5) 用线连接所有匹配的关键点。

3. 实验结果与分析

由实验描述可知，在实验中，阈值的选取会对检测的成功率其关键性的作用，太大的话会造成错配率高，而太小的话会造成漏检的情况，所以预知的设定尤为重要。我们参考相关试验参数和多次实验结果，最终将本文实验所用阈值选定为： $T = 0.6$ 。

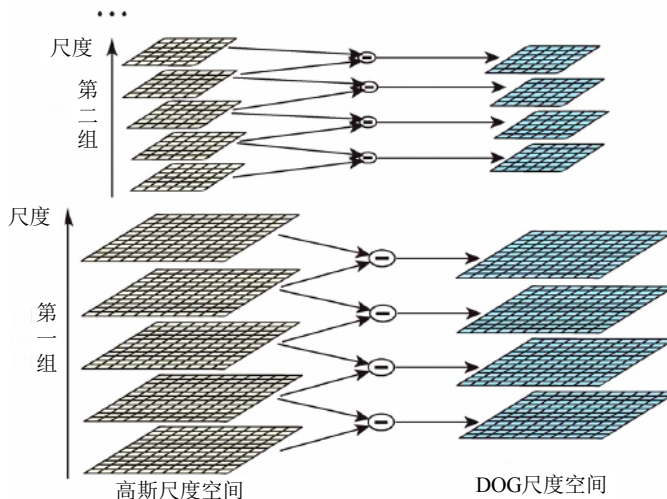


Figure 1. Differential Gaussian pyramid image
图 1. 差分高斯金字塔图像的构建

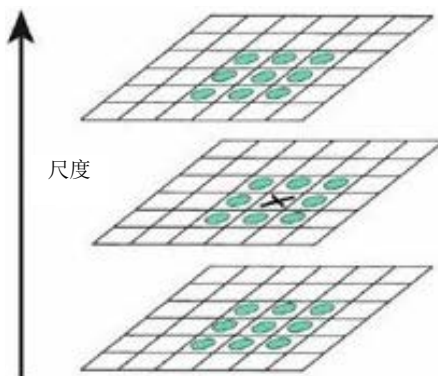


Figure 2. Extreme point detection
图 2. 极值点检测

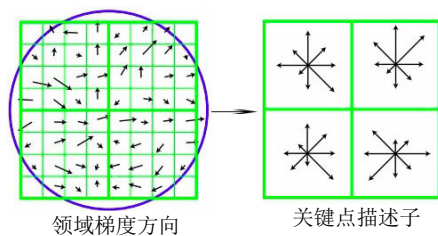


Figure 3. The construction of descriptors
图 3. 关键点描述子的构造

3.1. 复制变换移动粘贴篡改检测

为了准确评价本文算法的检测性能，从网上下载了一些内容各异的图片，然后利用 Photoshop 软件对图片进行了复制变换移动粘贴篡改的各种处理。本文试验是在 Windows xp, Matlab2010a 的环境下进行。由于篇幅限制，下面选取部分实验结果罗列。

首先对图像分别做普通平移复制粘贴篡改和复制粘贴篡改后加滤波修饰，然后分别利用 copy-move 算法和本文算法对篡改图像进行检测。得到检测结果如图 4 和图 5 所示。

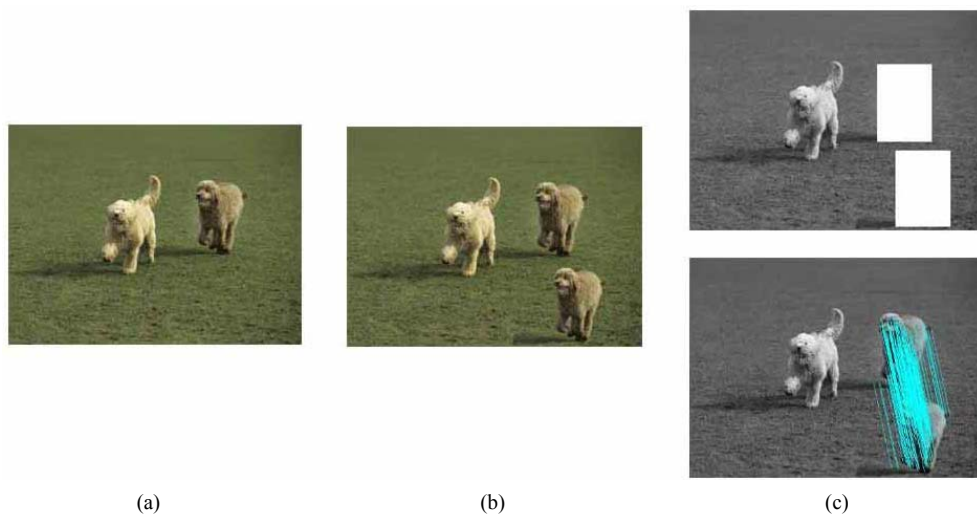


Figure 4. Test results in the case of translation. (a) Original image; (b) Tampered image; (c) Detected result

图 4. 平移复制粘贴实验结果。(a) 原始图像；(b) 篡改图像；(c) 检测结果

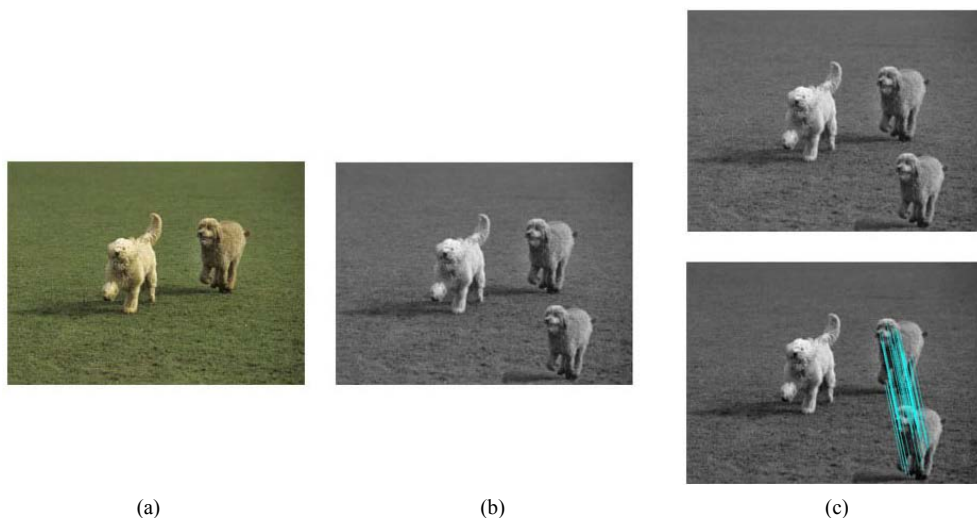


Figure 5. Test results in the case of Gauss filter ($\sigma = 1.6$). (a) Original image; (b) Tampered image; (c) Detected result

图 5. 高斯滤波($\sigma = 1.6$)。(a) 原始图像；(b) 篡改图像；(c) 检测结果

如图 1 和图 2 所示, 针对平移复制粘贴篡改, copy-move 算法和本文算法都能得到很好的实验结果, 但是对于经过后续篡改加滤波后, 本文算法仍然能够检测出结果, 而此时 copy-move 算法已经失去作用, 没办法得到实验结果。实验结果表明本文算法相对于传统算法在抗后续篡改方面具有一定的优势。下面对图像进行几何变换后复制粘贴进行试验, 分别对图像进行旋转(图 6 顺时针旋转 90 度)、缩放(图 7 缩放 1/2)、旋转 + 缩放(图 8 顺时针旋转 90 度, 缩放 1/2)等后续几何变篡改。而由于 copy-move 算法是通过将图像进行分块, 把每一个图像块当做一个模板, 然后遍历图像剩余的部分, 看是否存在跟模板完全一样的图像块来进行检测, 因此对于经过后续几何变换的篡改, 该算法完全失效, 而此时本文算法仍能得到较好的效果。下面将本文算法的实验结果罗列如下。

从图像的检测结果可以看出, 针对图像的普通复制粘贴篡改, 本文算法和 copy-move 算法都可以得

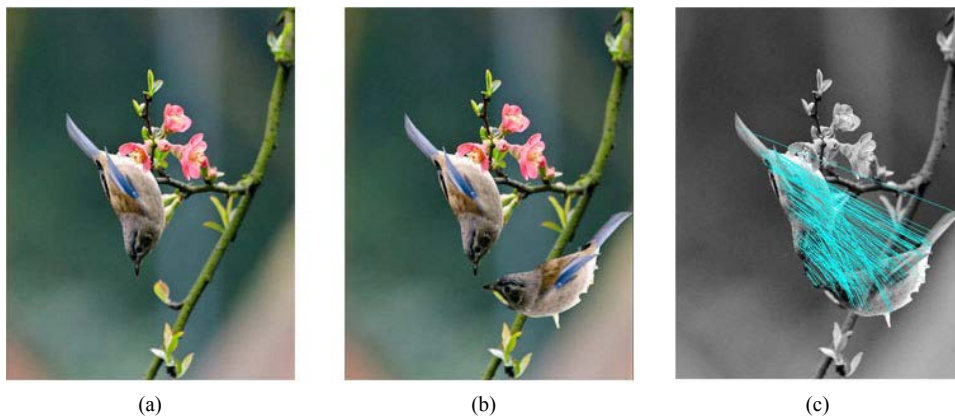


Figure 6. Test results in the case of rotation. (a) Original image; (b) Tampered image; (c) Detected result

图 6. 旋转后复制粘贴实验结果。(a) 原始图像；(b) 篡改图像；(c) 检测结果



Figure 7. Test results in the case of scaling. (a) Original image; (b) Tampered image; (c) Detected result

图 7. 缩放后复制粘贴实验结果。(a) 原始图像；(b) 篡改图像；(c) 检测结果

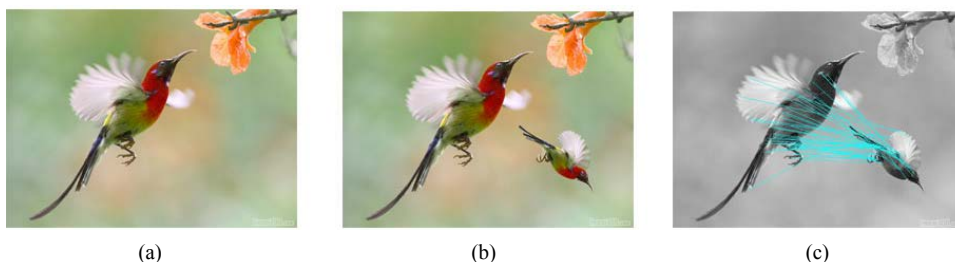


Figure 8. Test results in the case of rotation and scaling. (a) Original image; (b) Tampered image; (c) Detected result

图 8. 旋转 + 缩放后复制粘贴实验结果。(a) 原始图像；(b) 篡改图像；(c) 检测结果

到很好地检测效果，但是对经过后续修饰的篡改图像，copy-move 算法就完全失去了作用，而不管是简单的复制粘贴篡改还是经过后续修饰(滤波、旋转、缩放、旋转+缩放)的篡改图像，经本文算法检测后，图像的复制区域和经过几何变换的篡改区域之间被准确的对应起来。而传统的 copy-move 算法只能检测出普通复制粘贴的情况。由此可验证基于 SIFT 关键点的复制粘贴篡改检测算法对复制粘贴篡改中几何变换的篡改检测是有效可行的。

3.2. 算法比较

将本文算法和 copy-move 算法进行比较，从检测效果来进行比较。从网络上选取 20 幅图像，分别进行简单复制粘贴，旋转后复制粘贴，缩放后复制粘贴，组合变换(旋转+缩放)后复制粘贴变换，得到一个

Table 1. Comparison of detection results
表 1. 算法检测效果比较

	Copy-move 算法	本文算法
普通复制	√	√
普通复制(滤波)		√
旋转		√
缩放		√
旋转 + 缩放		√

包含 100 幅图像的图像库，分别利用 copy-move 算法和本文算法对图像库中图像进行检测。

将本文算法与 copy-move 算法进行比较，结果如表 1 所示，第一行表示不同算法，第一列表示不同的几何变换。由表中可以看出针对增加后续几何变换的篡改，copy-move 算法是失效的，而本文的基于 SIFT 的算法则可以得到很好地效果，不仅能够检测出简单复制粘贴的篡改，而且可以检测出经过后续旋转和缩放以及二者组合的篡改。

4. 结束语

针对传统的 copy-move 算法并不能检测出经过高斯滤波及几何变换等方式的复制粘贴篡改的不足，本文提出了一种基于特征点的复制粘贴算法，可以实现对高斯滤波、几何变换的检测。本文通过首先提取图像的特征点，然后利用欧氏距离进行特征点的匹配，最后通过自然图像中一般不会存在互相匹配的特征点的原理将相匹配的特征点用线连接起来，来检测出经过几何变换的篡改区域。经过试验证明此算法相对于传统的算法的不足有了极大的改进。实验结果表明，算法不仅可以检测出伪造图像中经平移变换的篡改区域，还可以检测出经旋转、缩放以及两者组合的几何变换。与其他的区域复制篡改算法相比，算法对伪造图像的检测有了较大的改进。

基金项目

国家自然科学基金项目(61070207, 61370195)，北京市自然科学基金项目(4132060)和“十二五”国家密码发展基金密码理论课题项目(MMJJ201201002)。

参考文献 (References)

- [1] 顾伟, 吕皖丽, 罗斌 (2009) 基于图像分类的矢量量化数字水印算法. *计算机应用研究*, 7, 2738-2740. <http://dx.doi.org/10.3969/j.issn.1001-3695.2009.07.097>
- [2] Fridrich, J., Soukal, D. and Lukáš, J. (2003) Detection of copy-move forgery in digital images. *Proceedings of Digital Forensic Research Workshop*.
- [3] Popescu, A.C. and Farid, H. (2004) Exposing digital forgeries by detecting duplicated image regions. *Comput.sci.dartmouth College Private Ivy League Res.univ*, 646.
- [4] Li, G.H., Wu, Q., Tu, D. and Sun, S.J. (2007) A sorted neighborhood approach for detecting duplicated image regions in image forgeries based on DWT and SVD. *Processing of 2007 IEEE International Conference on Multimedia and Expo*, Beijing, 2-5 July 2007, 1750-1753.
- [5] 王鑫, 轩波, 彭思龙 (2010) 基于融合的高分辨率彩色图像拷贝 - 变换 - 移动篡改检测. *中国图象图形学报*, 7, 1047-1053.
- [6] 骆伟祺, 黄继武, 丘国平 (2007) 鲁棒的区域复制图像篡改检测技术. *计算机学报*, 11, 1998-2007. <http://dx.doi.org/10.3321/j.issn:0254-4164.2007.11.013>
- [7] Lowe, D.G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 91-110. <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>

-
- [8] Amerini, I. (2011) A SIFT-based forensic method for copy-move attack detection and transformation recovery. *IEEE Transactions on Information Forensics & Security*, **6**, 1099-1110. <http://dx.doi.org/10.1109/TIFS.2011.2129512>
- [9] Dellinger, F., Delon, J., Gousseau, Y., et al. (2015) SAR-SIFT: A SIFT-like algorithm for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **53**, 453-466.
- [10] Zhang, J., Sang, H.S. and Shen, X.B. (2012) Improved SIFT matching algorithm with adaptive matching direction and scale restriction. *International Journal of Digital Content Technology and Its Applications*, **6**, 851-858.
- [11] 邹承明, 徐泽前, 薛栋 (2015) 一种基于分块匹配的 SIFT 算法. *计算机科学*, **4**, 311-315.
- [12] 周丽芬 (2014) 基于 SIFT 的图像匹配算法. *计算机与现代化*, **7**, 63-67.