

文章编号:1009-3087(2014)04-0075-06

图像模糊篡改盲取证

琚生根,夏欣,邵寿平,王能
(四川大学 计算机学院,四川 成都 610065)

摘要:为了提高模糊润饰篡改取证技术的准确率和召回率,提出新的盲取证技术。采用能量比和信息熵构建模糊边缘属性约简规则来检测源图像自然强边缘;设计保持边缘平滑的滤波结构增强源图像模糊区域,对滤波图像进行全边缘检测;设计人工模糊隶属度函数,计算源图像人工模糊隶属度集合;对所得边缘图像进行融合,利用人工模糊隶属度集合对融合图像进行处理,得到准确清晰的人工模糊边缘。实验结果显示,本文算法盲取证准确率和召回率较高,能较好区别人工模糊和自然模糊,篡改区域定位效果较好。

关键词:模糊润饰;属性约简规则;自然强边缘;边缘平滑滤波结构;人工模糊隶属度
中图分类号:TP391.41 **文献标志码:**A

Image Fuzzy Tampering with Blind Forensics

JU Shenggen, XIA Xin, SHAO Shouping, WANG Neng
(College of Computer Sci., Sichuan Univ., Chengdu 610065, China)

Abstract: A new forensics technology was proposed in order to improve the accuracy and recall rate vague of evidence tampering retouching techniques. By detecting the origin image's naturally strong edges on attribute reduction rules built by energy ratio and information entropy, using edge preserving smoothing filter structure to enhanced fuzzy region and detect the image's edges, designing artificial fuzzy membership function to compute every pixel's fuzzy membership and use it to detect the origin image's fuzzy edges, merging all the images and processing them through artificial collection of fuzzy membership, the accurately and clearly artificial fuzzy edges were obtained. Experimental results showed that the new algorithm gets a higher precision and recall rate, distinguishes artificial and natural blur better, and detects tampered region better.

Key words: fuzzy retouching; attribute reduction rules; naturally strong edge; edge smoothing filter structure; artificial fuzzy membership

随着图像编辑软件快速发展,数字图像越来越容易被篡改同时又不留下人眼可以观察的痕迹。这些篡改多为生活娱乐,但少部分会给如司法机关、学术界等领域带来危害^[1-2]。因此用于解决数字图像真实性问题的图像取证技术越来越受到人们重视。

依据是否有添加附加信息,图像取证技术可分为2类:主动取证和盲取证技术。主动取证技术^[3]主要通过检测图像中携带的数字水印和签名信息的完整性来判断图像的篡改与否。但该方法具有很大的局限性,随着信息的发展,通过图像本身信息直接判断图像篡改与否的图像盲取证技术的应用越来越

广泛,处于图像研究中的新兴前沿领域。

最简单的图像篡改方式是区域或者对象间的复制粘贴篡改^[4]。针对同一副图像内的区域篡改, Fridrich等^[5]首次提出了复制-篡改的取证问题,并利用字典排序寻找相似区域,但算法时间复杂度高;文献^[6-7]根据图像中不存在完全相同的区域块,提出了图像块特征的取证算法,目的是对特征进行降维以减少字典排序的计算量;中山大学的骆伟祺等^[7]利用图像块均匀分割得到维数较低的小块特征来增加检测结果的稳健性。在此基础之上,周琳娜等^[8]提出基于SIFT匹配来取证同幅图像复制-

收稿日期:2013-12-26

基金项目:国家自然科学基金资助项目(81373239);四川省科技支撑计划项目(2012GZ0091);四川大学青年基金项目(2012SCU11033)

作者简介:琚生根(1970—),男,副教授,博士。研究方向:图像篡改取证、医学信息处理等。

http://jsuese.scu.edu.cn

粘贴篡改的算法,比前面算法具有更快的速度和鲁棒性。Dybala 等^[9]将高阶微分图像中的复制-移动检测运用到区域篡改取证,寻找高阶微分克隆造成的相似区域。

对象篡改^[9]主要指在不同图像之间进行复制-拼接篡改,该篡改常常将复制对象的边界进行局部模糊处理,从视觉上达到掩盖篡改的目地,根据这一特点陈春宁等^[10]利用同态滤波实现图像对比图的增强;周琳娜等^[9]提出一种保持边缘平滑边缘结构增强模糊区域,能够很好地增强源图像的人工模糊边缘,但图像中存在的自然强边缘对实验结果造成很大的干扰;赵俊红^[11]利用粗糙集理论构造人工模糊边缘的约减属性和规则,能够很好地检测人工模糊区域并降低自然边缘区域对结果的干扰,但所检测的模糊边缘存在较多断点。

首先利用能量比和信息熵检测源图像自然强边缘,得到强边缘,并设计保持边缘平滑滤波结构来增强模糊区域,用 sobel 算子提取处理后图像的边缘,得到源全边缘。最后利用模糊隶属度函数处理源图像得到源图像像素的隶属度集合,融合全边缘图像和强边缘图像,用隶属度集合对结果进行人工模糊边缘提取,并用开运算处理结果图。实验显示本算法能够准确得定位篡改区域,较好地区分人工模糊和自然模糊,篡改边缘具有较好的连贯性。

1 自然强边缘检测

基于粗糙集理论^[12],利用能量比和信息熵构建判定图像区域为平滑区域或边缘区域的约简规则。

首先对源图像进行分块,大小为 8×8 , 计算每一个块的 DCT^[13] 系数,利用 DCT 系数计算图像块的能量比和信息熵。

能量比 R_h :

$$R_h = \frac{n-r}{r} \sqrt{\frac{\sum_{i=1}^r y_i^2}{\sum_{i=r+1}^n y_i^2}} \quad (1)$$

信息熵 H_f :

$$H_f = - \sum_{i=1}^n \left(\frac{|y_i|}{S} \lg \frac{|y_i|}{S} \right) \quad (2)$$

其中:

$$S = \sum_{i=1}^n |y_i| \quad (3)$$

式(1)中, r 表示低频像素的数目,对 8×8 图像块中, r 的理想取值是 10。能量比表示低频信息与高频信息比,值越少,细节成份越多;信息熵反映图像块能量集中的程度,值越大,说明所含细节和边缘的成

份越多。

根据粗糙集理论的属性约简步骤构建五条属性约简规则来提取源图像的自然强边缘:

规则 1: 当能量比大于某个很大的阈值 a 并且方差小于一个较小的阈值 b , 该区域为平滑区域;

规则 2: 当系数熵小于一个非常小的阈值 c 且方差小于一个较小的阈值 d , 该区域为平滑区域;

规则 3: 当能量比小于一个非常小的阈值 e , 该区域为边缘区域;

规则 4: 当系数熵大于一个非常大的阈值 f , 该区域为边缘区域;

规则 5: 当能量比小于一个较小的阈值 g 并且系数熵大于一个较大的阈值 h , 该区域为边缘区域。

规则中所有阈值均为经验阈值,通过改变不同的阈值,对图像库中的 50 幅图像进行的实验效果对比显示,得出算法理想阈值,程序中阈值的取值情况为 $a = 2.8, b = 1000, c = 2.5, d = 1000, e = 1.65, f = 3.92, g = 2.09, h = 3.90$ 。最后利用制定的规则对源图像进行强边缘检测,检测结果见图 1。

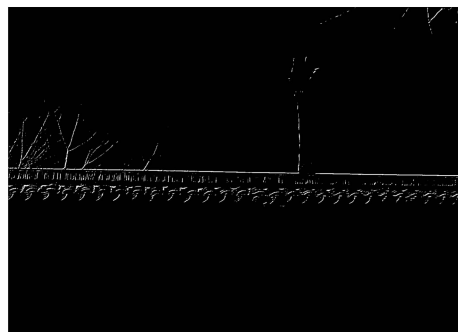


图 1 源图像自然强边缘

Fig. 1 Origin image's naturally strong edges

2 模糊边缘增强

人工模糊的特点是图像局部模糊区域内的像素值趋于平滑,但对模糊半径外的图像像素不做任何处理^[14]。模糊的本质是计算区域内像素灰度均值并给中心像素赋值^[15],若要消除这种由均值带来的模糊效应,需要在周围寻找不含边缘特性的像素点来代替该点的像素。设计一种保持边缘平滑滤波结构对源图像进行滤波,达到在保持原边缘的基础上,对模糊边缘进行增强的目的。

滤波公式如下:

$$\sigma_k = \sum_{f_j \in \varphi_k} |\tilde{f}_j - \tilde{f}_k|^2, \quad (4)$$

$$F(i, j) = \frac{\tilde{f}_k}{\sigma_k = \min(\sigma_1, \sigma_2, \dots, \sigma_9)}$$

式中: σ_k 表示第 k 个滤波结构中所有像素的方差; $\bar{f}_k$ 表示第 k 个滤波结构中素像素的均值; $F(i, j)$ 表示当前像素的新值, 数值为所有 9 个滤波结构中方差最小结构的像素均值。

为减少空间距离远的像素对结构中心像素的影响, 对每一个滤波结构中为中心像素欧式距离最远点进行剔除, 加强图像像素的局部相关性, 滤波结构最终设计如图 2 所示。

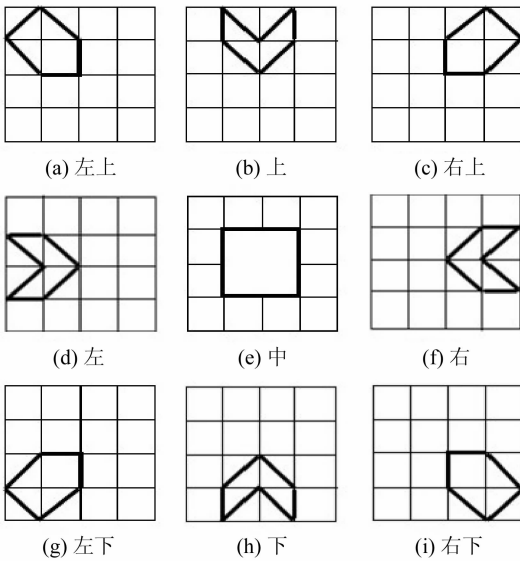


图2 保持边缘平滑滤波结构

Fig.2 Edge preserving smoothing filter structure

滤波结构共有 9 个子结构, 代表了中心像素周围的 9 个方向。分别求取每个滤波结构内像素的灰度方差, 再用方差最小的结构的像素均值赋值给模板中心像素。此操作能够对像素起到一个简单聚类作用, 对边缘进一步起到锐化的作用。

对滤波结果进行边缘检测, 结果如图 3 所示。



图3 模糊边缘增强图像

Fig.3 Fuzzy edge enhanced image

3 模糊隶属度集合

融合全边缘和强边缘图像, 得到模糊边缘, 需要剔除融合图像中的离焦模糊。人工模糊介于离焦模

糊和清晰边缘之间, 采用模糊集^[16]理论知识为像素赋予对应的人工模糊隶属度, 规则是人工模糊像素的隶属度最大, 平滑区域和自然清晰边缘趋于 0, 自然模糊处于中间。

隶属度函数需要分别的函数段包括平滑区域、清晰边缘、人工模糊边缘和自然模糊边缘; 平滑区域和清晰边缘的隶属度为最小值 0, 人工模糊边缘的隶属度为最大值 1, 自然模糊隶属度在区间 (0, 1) 中; 为了便于自然模糊隶属度值向人工模糊迭代, 自然模糊段函数设置为 2 次函数; 根据模糊半径的不同, 人工模糊函数段居中; 离焦模糊半径不同, 分布整个数域, 函数段处于第 2 和第 4 段, 综上所述, 所设计人工模糊隶属度函数如下:

$$\mu_{ij} = \begin{cases} 0, x(i, j) \leq a - b; \\ -\frac{(x(i, j) - a + c)^2}{(b + c)^2} + 1, a - b < x(i, j) < a - c; \\ 1, a - c \leq x(i, j) \leq a + c; \\ -\frac{(x(i, j) - a - c)^2}{(b + c)^2} + 1, a + c < x(i, j) < a + b; \\ 0, a + b \leq x(i, j) \leq \max \end{cases} \quad (5)$$

其中, $x(i, j)$ 表达式如下:

$$x(i, j) = |f(i, j) - \bar{F}| \quad (6)$$

式中 $f(i, j)$ 表示灰度值, $\bar{F}$ 表示模板内像素均值, $x(i, j)$ 表示当前像素会模板内像素均值的差值绝对值。 a, b, c 为抛物线函数参数, 由离焦半径、模糊半径和对称轴决定, 取参数为 $a = 4, b = 3, c = 1$ 。式 (4) 物理意义为:

1) 差值绝对值很小时, 表明是当前区域为平滑区域; 差值绝对值相差很大时, 表明像素处于边缘区域或者属于噪声点, 所以为像素的人工模糊隶属度赋值为最小值 0;

2) 差值绝对值在区间 $(a - c, a + c)$ 时, 表明像素处于具有一定半径的模糊区域内, 属于人工模糊边缘, 像素的模糊隶属度赋值为最大值 1;

3) 差值绝对值在 $(a - b, a - c)$ 和 $(a + c, a + b)$ 时, 说明像素属于离焦模糊, 根据表达式为该像素的隶属度赋值, 大小在区间 (0, 1) 中。

隶属度函数图像如图 4 所示。

利用所设计的隶属度函数对源图像进行处理, 获取源图像每个相似点的人工模糊隶属度集合, 像素均值以 8×8 邻域内所有像素进行计算。对应于图像的人工模糊隶属度集合 X 表达式如下:

$$X = \bigcup_{i=1}^M \bigcup_{j=1}^N \mu_{ij} \times f(i, j) \quad (7)$$

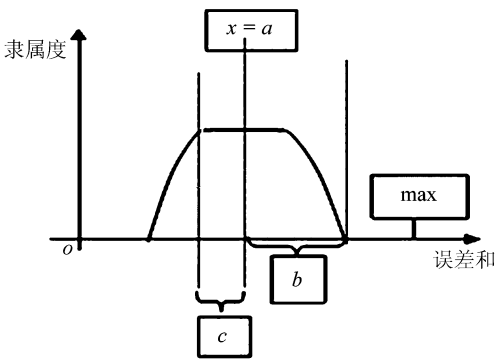


图 4 模糊隶属度函数图形

Fig.4 Image of the fuzzy membership function
其中: $i = 1, 2, 3, \dots, M; j = 1, 2, 3, \dots, N; \mu_{ij}$ 表示人工模糊隶属度; $f(i, j)$ 表示像素灰度值。

4 图像模糊篡改取证

4.1 算法流程

本文算法流程图如图 5 所示。

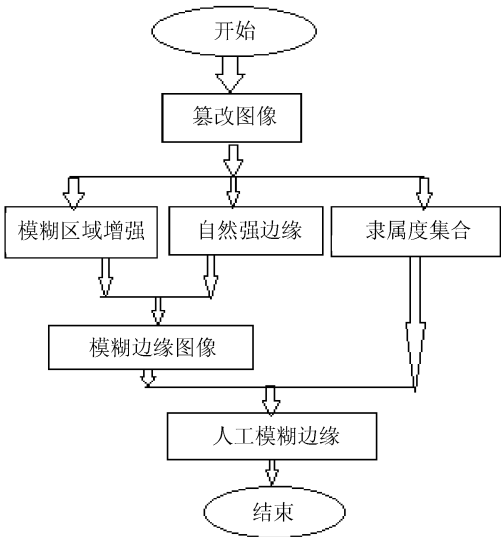


图 5 算法流程图

Fig.5 Algorithm flowchart

具体步骤如下:

- 1) 根据粗糙集理论和属性约减步骤,得到检测自然强边缘的属性约减规则,并利用这 5 条规则对源图像的灰度图像进行边缘检测,得到自然强边缘。
- 2) 求取源图像的灰度图像,用设计的保持边缘平滑滤波结构对图像进行滤波处理,并用 sobel 算子检测模糊区域增强后的图像,得到源图像的全边缘。
- 3) 根据模糊集的理论知识,设计人工模糊隶属度函数,对源图像灰度图像进行一一处理,得到源图像与各个像素相对应的人工模糊隶属度集合。
- 4) 融合全边缘图像和自然强边缘图像,去除全

边缘中的强边缘,然后利用第 3) 步所得的隶属度集合对模糊边缘进行处理,处理的规则是用模糊边缘图像各像素点与相应的隶属度进行乘积,并设定一定的阈值,大于阈值设置为 255,小于阈值的像素值设置为 0;最后利用闭运算对结果进行处理。

取证结果如图 6 所示。

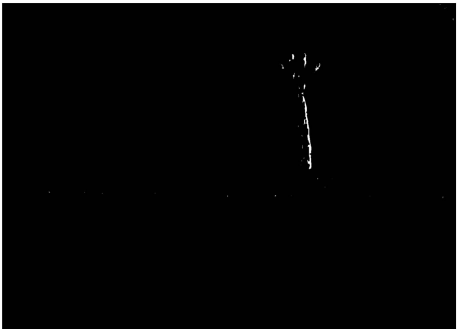


图 6 取证结果(人工模糊边缘)

Fig.6 Forensics result (artificial fuzzy edges)

4.2 实验数据

素材中背景图片以自然风景为主,纹理简单复杂参半(图 7)。

复制对象以动物为主,模糊润饰操作为羽化操作。羽化半径根据对象以及边缘情况变化而变化,从 1 到 10 不等,如图 8 所示。



图 7 源图像

Fig.7 Source images

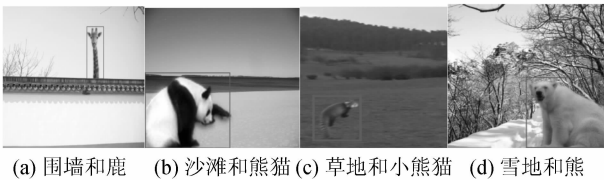


图 8 篡改图像

Fig.8 Tampered images

为了对取证效果的可行性进行分析,对周琳娜、赵俊红以及本文算法进行了结果比对,篡改结果如图 9 所示。

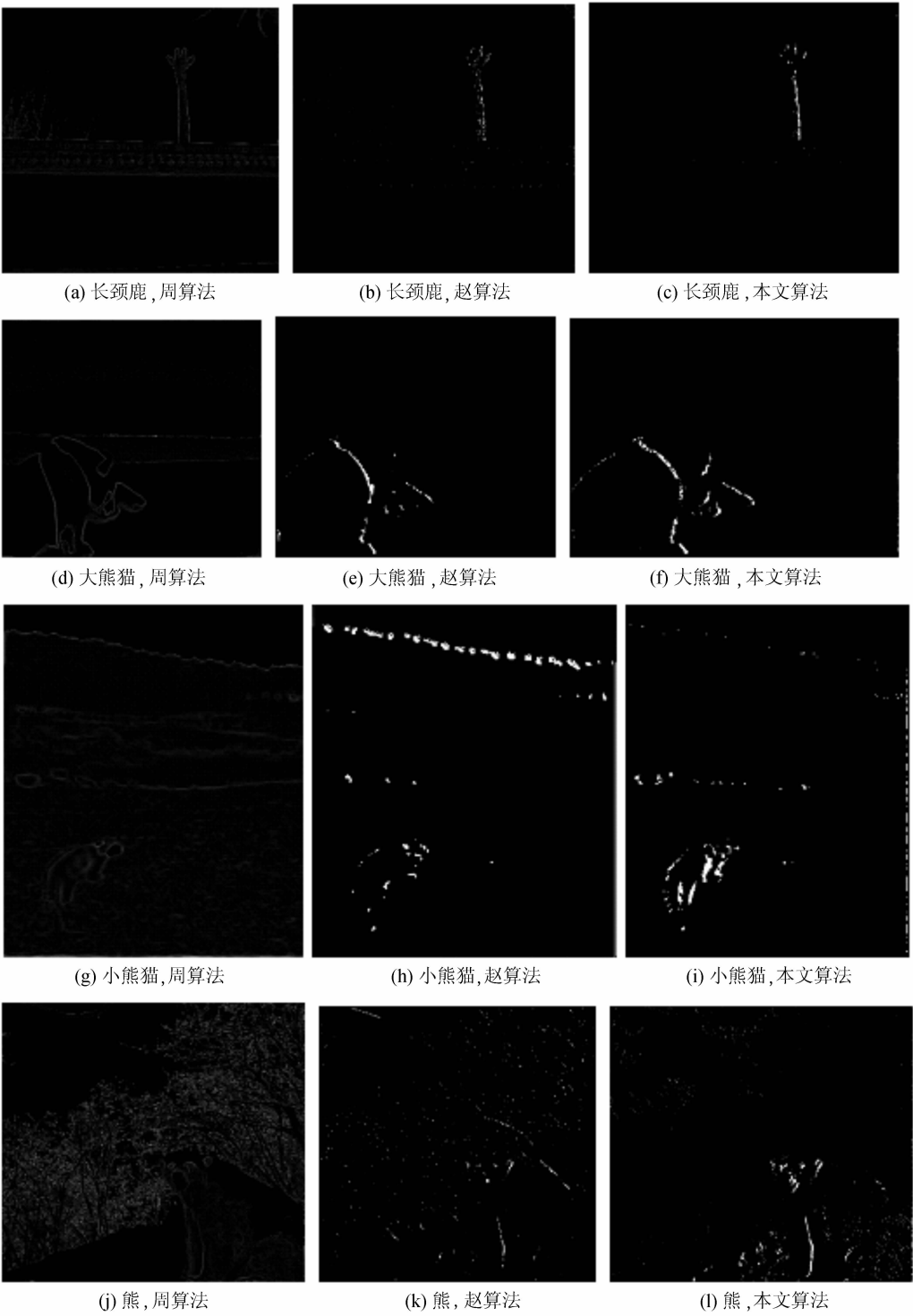


图 9 取证结果(周、赵、新算法)

Fig. 9 Results (Zhou,Zhao,and new algorithm)

4.3 结果与分析

取证准确率、召回率和 F1 度量对比如图 10 所示。

实验结果中,周的方法实验效果较差,不能够对人工模糊和自然模糊进行很好的区别,取证区域模糊;赵算法能够达到篡改取证目的,结果较准确,但

篡改区域像素的召回率较低,并且依然存在一些错误的取证区域,比如长颈鹿上的斑点,小熊猫的远山,以及熊中的零落树枝等,人工模糊边缘不连贯;本文算法取证区域清晰连贯,能够很好地区别人工模糊和自然模糊,具有较高的取证召回率,实验结果具有较强的说服力。

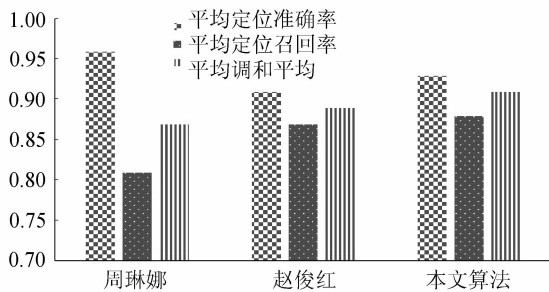


图10 周、赵、本文算法结果对比

Fig.10 Results contrast of Zhou,Zhao,and new algorithm

5 结 论

主要对篡改取证中的人工模糊篡改手段进行盲取证算法设计。首先利用能量比和信息熵提取图像的粗糙强边缘;然后设计保持边缘平滑滤波结构对图像进行模糊区域增强,得到图像全边缘图像;第3步通过粗糙集理论设计属性约束规则,得到源图像的模糊隶属度集合和模糊边缘;通过融合前3幅图像,得到最后的篡改区域。实验结果显示,本算法能很好地寻找人工模糊篡改区域,具有一定的应用价值。今后可以在提高算法运行速度上继续努力,让强边缘计算的时间复杂度尽量减少,同时对模糊隶属度函数进行进一步的改善,减少模糊边缘中的断点现象,同时尽量消除其他孤立的干扰点。

参考文献:

- [1] Zuo Juxian, Liu Benyong. Detection for typical tampering operations in a forged image[J]. Jouranl of Image and Graphics, 2012, 17(11): 1367 - 1375. [左菊仙, 刘本永. 伪造图像典型篡改操作的检测[J]. 中国图象图形学报, 2012, 17(11): 1367 - 1375.]
- [2] Fridrich J. Digital image forensics[J]. IEEE Signal Processing Magazine, 2009, 26(2): 26 - 37.
- [3] Fridrich J. Image watermarking for tamper detection[C]//Proceedings of International Conference on Image Processing. Chicago: IEEE, 1988: 404 - 408.
- [4] Farid H. Image forgery detection[J]. IEEE Signal Processing Magazine, 2009, 26(2): 16 - 25.
- [5] Fridrich J, Soukal D, Lukas J. Detection of copy-move forgery in digital images[C]//Proceedings of the Digital Forensic Research Workshop. Cleveland: OH, 2003: 5 - 8.
- [6] Wu Qiong, Li Guohui, Sun Shaojie, et al. Image forgery detection region based on wavelet and singular value decomposition[J]. Micro Computer System, 2008, 29(4): 730 - 733. [吴琼, 李国辉, 孙韶杰, 等. 基于小波和奇异值分解的图像复制伪造区域检测[J]. 小型微型计算机系统, 2008, 29(4): 730 - 733.]
- [7] Luo W Q, Huang J W, Qiu G P. Robust detection of region-duplication forgery in digital image[C]//Proceedings of the 18th International Conference on Pattern Recognition. Hong Kong: IEEE, 2006: 746 - 749.
- [8] 周琳娜, 张茹, 郭云彪. 数字图像内容取证[M]. 北京: 高等教育出版社, 2011.
- [9] Dybala B, Jennings B, Letscher D. Detecting filtered cloning in digital images[C]//Proceedings of the 9th Workshop on Multimedia and Security. Manchester: IEEE, 2007: 43 - 50.
- [10] Chen Chunling, Wang Tingjie. The use of image contrast enhancement filter Thai children in the frequency domain[J]. Microcomputer Information, 2007, 23(3): 264 - 266. [陈春宁, 王廷杰. 在频域中利用童泰滤波增强图像对比度[J]. 微计算机信息, 2007, 23(3): 264 - 266.]
- [11] Zhao Junhong. Passive forensics algorithm of tampered image by detecting edge blurring manipulation[J]. Journal of South China University of Technology: Natural Science Edition, 2011, 39(7): 77 - 82. [赵俊红. 篡改图像边缘模糊操作的被动取证算法[J]. 华南理工大学学报: 自然科学版, 2011, 39(7): 77 - 82.]
- [12] Qiang He, Congxin Wu, Degang Chen, Suyun Zhao. Fuzzy rough set based attribute reduction for information systems with fuzzy decisions[J]. Knowledge-Based Systems, 2011, 2(9): 689 - 696.
- [13] Huang Yanping, Lu Wei, Sun Wei, et al. Improved DCT-based detection of copy-move forgery in images[J]. Forensic Science International, 2011, 8(1): 178 - 184.
- [14] Bao Hua, Fan Yubo, Rao Changhui, et al. Fast median filtering algorithm based on the average lookup[J]. Journal of Sichuan University: Engineering Sciences Edition, 2011, 43(2): 76 - 79. [鲍华, 樊瑜波, 饶长辉, 等. 基于均值查找的快速中值滤波算法[J]. 四川大学学报: 工程科学版, 2011, 43(2): 76 - 79.]
- [15] Popescu A C, Farid H. Exposing digital forgeries by detecting traces of resampling[J]. IEEE Transactions on Signal Processing, 2005, 53(2): 758 - 767.
- [16] Verstraete J, Van DerCruyssen B, De Caluwe R. Assigning membership degrees to points of fuzzy boundaries[J]. IEEE Transactions on Fuzzy Information Processing Society, 2000, 13(15): 444 - 447. (编辑 杨 蓓)