

基于数字图像边缘特性的形态学滤波取证技术

周琳娜^{1,2}, 王东明³, 郭云彪², 杨义先¹

(1. 北京邮电大学信息安全中心, 北京 100876; 2. 北京电子技术应用研究所, 北京 100091;

3. 中国证券监督管理委员会信息中心, 北京 100032)

摘 要: 针对数字伪造合成图像中最常用的模糊操作, 本文提出了一种基于图像形态学滤波边缘特性的盲取证算法; 该方法用同态滤波和形态学滤波的方法增强了模糊操作的图像边缘, 分析了离焦模糊和人工模糊边界特点, 利用离焦模糊和人工模糊的边缘特性实现了针对数字图像伪造的被动盲取证; 实验表明, 该方法能够有效地检测伪造数字图像的模糊操作痕迹, 并能对已模糊操作的图像局部进行准确定位。

关键词: 数字取证; 图像处理; 形态学滤波; 模糊操作; 同态滤波

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2008) 06-1047-05

Exposing Digital Forgeries by Detecting Image Blurred Mathematical Morphology Edge

ZHOU Lin^{1,2}, WANG Dongming³, GUO Yunbiao², YANG Yixian¹

(1. Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Beijing Application Institute of Electronic Technology, Beijing 100091, China;

3. China Securities Regulatory Commission Information Center, Beijing 100032, China)

Abstract: A novel image forensic detection scheme for digital images was proposed in this paper, which can determine whether an image is tampered or not using blur operation. The difference between artificial blur and defocus blur can be indicated by edge processing using homostasis filtering and mathematical morphology. The blue tampered regions can be detected by blur edge analysis. Experimental results demonstrate the effectiveness of the proposed scheme.

Key words: digital forensics; image processing; mathematical morphology; blur; homostasis filtering

1 引言

随着图像数字化摄取设备的普及和图像编辑软件功能(如 Adobe PhotoShop)的多样化, 数字图片的伪造和篡改在今天已经是一件很容易的事情。有关统计数据表明, 截止到 2005 年 10 月, 仅 PhotoShop 高手们在 www.worth1000.com 上公开的伪造数字图像就有 178, 582 张。有些伪造的数字图像只是出于艺术角度, 也有仅仅为个人使用, 但是有些伪造的数字图像却会对社会造成很大影响, 如 2003 年洛杉矶时报刊登的伊拉克战争图, 2004 年网上流传的美国参议员 John Kerry 和女演员 Jane Fonda 同台反对越南战争图^[1]等等。证明这些图片的来源与真实性, 均离不开数字图像取证技术, 所以, 作为证明数字图像完整性和原始性的数字图像取证技术已成为计算机取证技术的一个新的热点。

数字图像取证是计算机取证的一个分支, 是对源于数字图像资源的数字证据进行确定、收集、识别、分析及出示法庭的过程。不同于原始的计算机取证, 数字图像取证主要是研究图像的完整性和原始性。先前的数字图像取证主要是基于数字水印的主动认证, 如易碎数字水印技术和数字签名技术提供了很多验证数字篡改的方法, 但这些方法都是通过提前向原始数字照片中嵌入验证信息实现的, 属于主动取证。随着图像篡改技术迅速发展, 主动取证技术由于受到应用条件的限制, 已经无法从根本上遏制图像篡改, 现在数字图像取证更注重被动取证技术研究。

数字图像取证是针对于图像篡改的技术, 要进行图像取证需对常用的图像篡改手段清楚的了解。Hany Farid 在文献[2]中把常用的图像篡改技术分为 6 类, 即合成图像、变体图像、润饰图像、增强图像、计算机生成

图像以及绘画图像,其中最常用的是合成图像,即指由两幅以上的数字图像,通过剪切各自的一部分合并生成的图像.大多数利用合成图像进行的图像篡改,都需进行拼接、粘贴等操作,为了消除在拼接边缘产生的视觉或统计上的畸变,都会采用模糊、淡化、渐变等润饰操作进行后期处理以消除拼接留下的人眼可察觉的伪造痕迹.因此,对图像合成和润饰的检测是数字图像取证研究的重点.文献[3~10]提出了针对简单图像拼接的合成数字图像取证算法,但是这些算法都无法对拼接后进行过润饰处理的数字图像进行检测.本文针对数字伪造合成图像中最常用的模糊操作,用同态滤波和形态学滤波的方法提取数字图像的边缘,提出了一种基于图像形态学滤波边缘特性的盲取证算法.实验表明,该方法能够有效地检测出数字图像中模糊操作的痕迹,并能对图像拼接的边界进行准确定位.

2 模糊操作与图像边缘特性

为了消除图像伪造在拼接边缘产生的视觉或统计上的畸变,伪造者通常会在图像拼接后采用模糊、淡化、渐变等润饰操作以消除简单拼接留下的伪造痕迹.其中,模糊操作是最常用的手法之一.模糊操作的基本原理是对图像的局部邻近像素值进行邻域灰度平均,其目的是为了将图像伪造所造成的拼接边缘淡化甚至消除.

2.1 人工模糊操作的数学表达

对图像进行模糊操作时,需指定模糊模式,用户可以选择不同的模糊半径和强度对图像进行模糊操作.从 Photoshop 软件实现的算法看,模糊操作就是用移动平均滤波对选定的图像区域进行运算而产生平滑的结果,不同的模糊模式可以通过选择不同的移动平均滤波函数和滤波窗的大小来实现,模糊强度的高低则反映在模糊方式的选择上.模糊操作可以看作是将输入图像中的像素 (i, j) 的邻域平均灰度确定为输出图像像素 (i, j) 的值,相当于利用所有元素都用 $1/n^2$ 的加权矩阵进行空间滤波,当使用 $n \times n$ 的正方形模糊邻域时,模糊操作可表示为式(1):

$$g(i, j) = \frac{1}{n^2} \sum_{k=-[n/2]}^{[n/2]} \sum_{l=-[n/2]}^{[n/2]} f(i+k, j+l) \quad (1)$$

其中, n 表示模糊邻域大小,即滤波器窗大小,它是一个非负整数;求灰度平均值的邻域 n 取得越大,模糊程度越强.人工模糊操作的结果在图像灰度级上的体现是压缩了图像选定模糊区域的灰度级.

2.1.2 图像离焦模糊和人工模糊

一些未经篡改的自然图像由于其场景范围广,会存在由于焦距不准而造成的自然模糊图像块.如果要对伪造篡改图像进行较精确的取证,就必须在技术上能够区分开自然模糊和人为模糊.

从人工模糊的数学表达式(1)可以看出,人工模糊是在选定模糊邻域内所有像素的统计平均,在模糊半径内的模糊滤波是等值处理的,而对于模糊半径以外的图像像素则不做任何处理.而相机离焦模糊由于不存在人工选定的模糊半径,其模糊滤波相对于聚焦点向外成射线状无级扩散,逐渐变小.这两种模糊效果的对比示意图如图1所示.

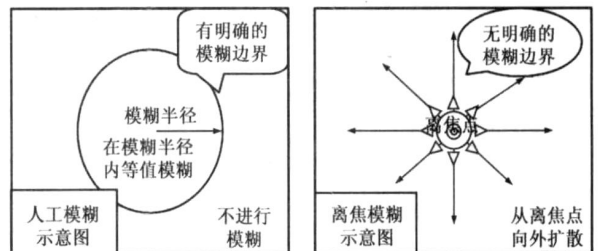


图1

由图1可以看出,人工模糊和相机离焦模糊的最大区别在于人工模糊有明确的模糊边界,在模糊边界内外的图像像素相对于模糊的概念区别较大,而相机离焦模糊没有模糊边界,较人工模糊会表现得更加均匀.如果可以找到合适的图像处理方法准确抵消掉人工模糊的邻域平均灰度滤波的模糊效果,对人工模糊操作中的模糊边界和边界内外像素的区别进行放大,就可以准确区分离焦模糊和人工模糊.通过分析和实验,发现同态滤波^[11]就是这样一种有效的处理算法.

2.1.3 同态滤波放大人工模糊边界

经过伪造粘贴模糊操作后的图像大都会处于这样一种状态:其拼接边缘由于经过了模糊处理,灰度级范围降低,即模糊的拼接边缘分不清物体的灰度层次和细节,而其余未经处理的正常边缘一般灰度级动态范围较大,即层次鲜明,黑白反差大.此时,若要对模糊处理后的拼接边缘进行加强放大,采用一般的线性变换难以奏效.因为若统一扩展图像灰度级虽可以提高物体图像的反差,但会使动态范围更大,同时也加强了未经处理部分的正常边缘,容易造成误判;而统一压缩图像的灰度级,则使人工模糊处理部分的灰度层次和细节更加不清楚,在取证时则无法设定阈值.针对这种情况,我们选用同态滤波的方法,把频率过滤和灰度变换结合起来,在频域中同时将图像亮度范围进行压缩和将图像对比度进行增强,达到既压缩整幅图像的灰度动态范围,又能让模糊的拼接边缘灰度级扩展,从而放大人工模糊操作中的模糊边界和边界内外像素的区别,提高取证的准确性.

同态滤波对图像不同部分的处理是不同的:图像的未经伪造部分对应的低频段照明分量,该部分图像清晰,灰度级范围大,但对我们的取证判断不能提供取证的有用信息,因此同态滤波过程中将其压缩;伪造图

像中拼接边缘被模糊的部分对应通过邻域平均灰度后的高频段反射分量, 该部分图像灰度级范围很小, 层次不清, 细节不明, 也正是在取证过程中应重点研究的信息, 为此在同态滤波增强过程中将其扩展, 以获得更多的信息. 图 2 为做同态滤波后的效果及取边缘对比图, 经同态滤波后的图像模糊处的边缘得到增强.

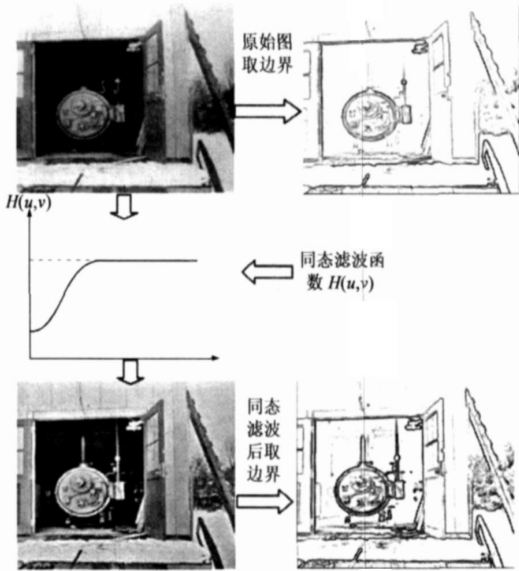


图 2

同态滤波器函数 $H(u, v)$ 以不同的方法影响傅里叶变换的高低频成分. 图 2 中显示了这种滤波器图, 其同态滤波器函数 $H(u, v)$ 表达式如式(2).

$$H(u, v) = (G_1 - G_2) [1 - e^{-\alpha D^2(u, v) / D_0^2}] + G_2 \quad (2)$$

其中, D_0 为滤波器的高通截止频率, $D(u, v)$ 是从点 (u, v) 到频率平面原点的距离, 即 $D^2(u, v) = u^2 + v^2$; 常数 c 被引入用来控制滤波器函数斜面的锐化, 它在 G_1 和 G_2 之间过渡. 且 $G_2 < 1, G_1 > 1$.

图像经过了同态滤波后, 放大增强了图像经人工模糊处理后的拼接边缘, 此时就可以设定合适的阈值, 采用数学形态学中的腐蚀运算, 排除图像的自然边缘, 提取定位出图像拼接伪造的区域.

3 利用形态学滤波的模糊操作检测算法

数学形态学以形态为基础对图像进行分析, 其基本思想是用具有一定形态的结构元素去量度和提取图像中的对应形状以达到对图像分析和识别的目的. 数学形态学图像处理有两种基本运算: 腐蚀与膨胀.

图像腐蚀是一种收缩变换, 其结果是把图像连接成分的边界点去掉从而把图像的边缘缩小一层的处理, 是将结构元素 B 在图像 A 中平移, 所有可以填入图像 A 内部的结构元素的点组成的新图像即为腐蚀的结果图像, 可表示为:

$$A \ominus B = \{X | B + X \subseteq A\} \quad (3)$$

图像膨胀就是把图像的连接边界扩大一层的处理, 是用结构元素 B 内的所有点作为参考点来平移输入图像 A , 然后计算平移结果的并集, 可表示为:

$$A \odot B = G\{A + X : X \subseteq B\} \quad (4)$$

数学形态学的运算主要针对二值图像, 用数学形态学滤波的方法对边缘二值图像的处理分析会较直接对灰度图像的处理分析更加精确. 因此, 在进行边缘检测时将灰度图像转为二值图像进行形态学运算处理可以减少图像灰度噪声对检测结果的干扰, 图 3 为用数学形态学的膨胀和腐蚀对边缘二值图像的处理结果.

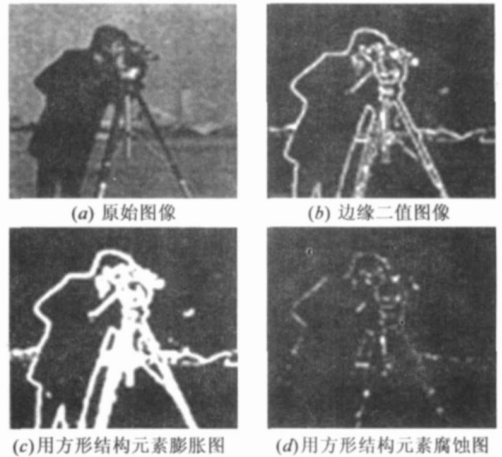


图 3

而人工模糊操作是用加窗滤波器函数对选定模糊半径内的图像像素进行滤波的结果, 其基本原理是对图像的模糊半径内的邻近像素值进行邻域平均灰度而产生平滑, 在数学形态学上可以看成对图像的局部边缘进行元素膨胀的结果. 图 4 为对一幅图像进行人工边缘模糊操作前后数学形态学滤波的边缘对比图.

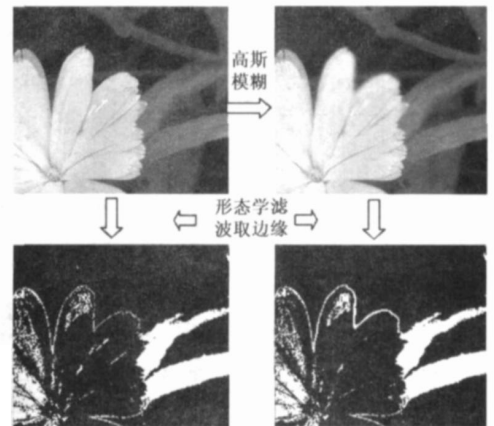


图 4

在用形态学滤波对边缘二值图像进行处理时, 构造合适的结构元素 (SE) 十分重要. 结构元素的形状、尺寸的选择是能否成功提取出有效边界信息的关键. 大的结

构元素可以去掉大的噪声颗粒,但也会腐蚀掉部分细节的拼接边缘,容易造成漏判;小的结构元素腐蚀不了层次宣明的部分正常边缘,容易造成误判;考虑到人工模糊的模糊半径大都是为三个像素左右,结构元素的尺寸选择为三像素的方形结构.经多次实验,选择 3×3 的方形结构:

$$SE = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (5)$$

通过对人工模糊的图像边缘特性分析可以看出,形态滤波学的腐蚀操作可以收缩掉部分图像/弱0边缘,保留下/强0边缘,对经过同态滤波修正增强后的二值边缘图像进行腐蚀操作,就可以收缩掉同态滤波未增强的图像正常区域/弱0边缘,同时保留同态滤波增强的图像篡改区域/强0边缘.

综上所述,对图像中模糊操作进行检测的过程如下:

(1) 设计合适的同态滤波函数,对预取证图像的边缘进行边缘灰度级扩展处理.在频域中同时将图像亮度范围进行压缩和将图像对比度进行增强,在对正常图像边缘影响较小的前提下,放大增强模糊处理后的边缘.

(2) 提取经同态滤波后图像的二值边缘图像信息.

(3) 选择合适的结构元素(SE)对二值边缘图像信息进行腐蚀运算,腐蚀收缩掉未经增强的图像正常边缘,提取经增强处理的图像模糊拼接边缘,从而定位图像的伪造区域.

4 仿真及实验结果

我们用 Matlab6.5^[12] 仿真了上述用数学形态学进行伪造图像模糊边缘算法中的同态滤波和腐蚀运算过程,并对用数码相机拍摄的原始图像和网络上的伪造篡改图像进行了检测取证,图 5 为对 www.worth1000.

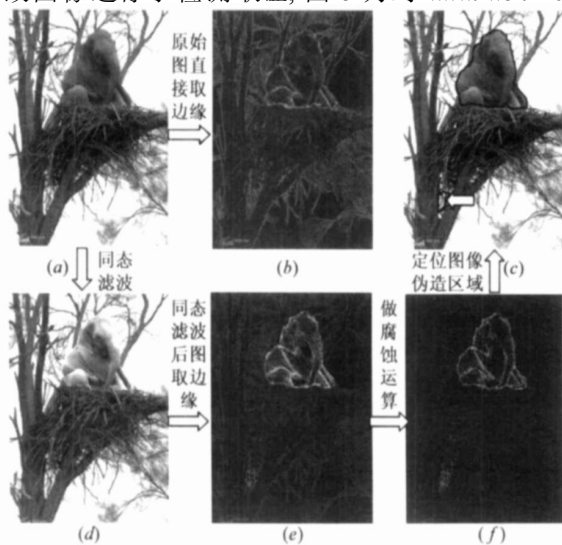


图 5

com 上的一幅伪造图片进行检测取证的过程:

从图 5(f) 中的黑色框定的区域为检测出的经过模糊操作的图像局部区域,其中有箭头所指黑框部分为错检部分.从检测结果中可以看出,经同态滤波处理后用数学形态学进行人工模糊边缘处理检测的取证方法能够较为准确地检测出经过模糊操作的伪造和篡改图像,并可以较精确地定位出图像的篡改伪造局部区域.

为了验证本算法对相机的离焦模糊和人工模糊的区分,我们用多种品牌数码相机拍摄的图像、用 Photo shop CS8.0.1 人工模糊图像及从 www.worth1000.com 上下载了部分篡改图像和原始图像进行了检测取证,结果如表 1:

表 1

| 图像来源 | 图像数 | 正确数 | 错误数 | 正确率(%) |
|------------------|-----|-----|-----|--------|
| Worth1000 上原始图像 | 42 | 41 | 1 | 97.6 |
| Worth1000 上合成图像 | 65 | 53 | 12 | 81.5 |
| Sony N1 原始图像 | 205 | 192 | 13 | 93.6 |
| Kodak DC290 原始图像 | 118 | 103 | 15 | 87.3 |
| Sony N1 人工模糊图像 | 51 | 48 | 3 | 94.1 |
| Sony N1 篡改合成图像 | 36 | 33 | 3 | 91.6 |
| NikonE5700 合成图像 | 25 | 22 | 3 | 88 |

5 总结与展望

本文从数学形态学的角度详细分析和探讨了模糊操作对数字图像边缘特性影响,提出了一种基于数学形态学的伪造图像模糊边缘提取和定位的方法.该方法针对模糊操作对数字图像边缘特性影响,设计了合适的同态滤波的方法增强了人工模糊的边缘特性,去除了自然模糊的干扰,然后采用数学形态学中的腐蚀运算,腐蚀收缩掉未经增强的图像正常边缘,提取定位出经增强处理的图像模糊拼接边缘.实验表明,该方法针对通过检测拼接和模糊制作的伪造篡改图像的边缘,可以准确检测和定位图像的拼接篡改.

数字图像篡改改造的方法多种多样,其对图像特性的影响也是多方面的,基于数字图像单一特性取证技术的检测取证精度随着伪造者伪造技术的提高而越发显得力度不足,因此,多角度、全方位地考虑数字图像篡改伪造对图像特性的影响,以提高数字图像取证的精确度以对抗更高水平的数字图像篡改改造,将是数字图像取证的下一步研究方向.

参考文献:

- [1] K Hafner. The camera never lies, but the software can [J]. In New York Times, 2004, 10(2): 145- 149.
- [2] Hany Farid. Creating and detecting doctored and virtual images [A]. Implications to The Child Pornography Prevention Act [C]. Dartmouth College, USA: Department of Computer Sci2

- ence, 2004. 280- 291.
- [3] J Luk Í, J Fridrich, M Goljan. Detecting digital image forgeries using sensor pattern noise[A]. In Proceedings of the SPIE[C]. California USA, 2006. 362- 372.
- [4] T T Ng, S F Chang, Q Sun. Blind detection of photomontage using higher order statistics[A]. IEEE International Symposium on Circuits and Systems[C]. Vancouver, Canada, 2004, 5: 688 - 691.
- [5] Fridrich J, Soukal D, Luk Í J. Detection of copy-move forgery in digital images[A]. In Proceedings of DFRWS[C]. Berlin: Springer-Verlag Press, 2003. 289- 302.
- [6] Ng T T, Chang S F. A model for image splicing[A]. In Proceedings of IEEE ISCAS[C]. New York: Academic Press, 2004. 1169- 1172
- [7] Lyu S. Natural image statistics for digital image forensics[R]. Hanover: Dartmouth College Department of Computer Science Technical, 2005. TR2005- 557.
- [8] A C Popescu, H Farid. Exposing digital forgeries in color filter array interpolated images[J]. IEEE Trans on Signal Processing, 2005, 53(2): 3948- 3959.
- [9] T T Ng, S F Chang. A data set of authentic and spliced image blocks[R]. Columbia: Columbia University, ADVENT Technical, 2004.
- [10] Popescu A C, Farid. Exposing digital forensics by Detecting duplicated image regions[R]. Hanover: Dartmouth College Computer Science Technical, 2004. TR2004- 515.
- [11] 龚声蓉. 数字图像处理与分析[M]. 北京: 清华大学出版社. 2006
Gong Shengrong. Digital Image Processing and Analysis[M]. Beijing: Tsinghua University Press, 2006. (in Chinese)
- [12] 孙兆林. MATLAB 6. X 图像处理[M]. 北京: 清华大学出版社. 2002.
Sun Zhaolin, MATLAB 6. X Image Processing[M]. Beijing: Tsinghua University Press, 2006. (in Chinese)

作者简介:



周琳娜 女, 1972 年 4 月出生于湖南省邵阳市, 1993 年毕业于解放军信息工程大学通信工程系, 现为北京电子技术应用研究所副研究员。主要研究方向为图像分析与处理、信息隐藏、多媒体信息安全。E-mail: linnazhou1972@hotmail.com

王东明 男, 1972 年 10 月出生于河北衡水, 1993 年毕业于解放军信息工程大学通信工程系, 现为中国证券监督管理委员会副研究员。主要研究方向为多媒体信息安全、网络安全。E-mail: wdmzln@263.net