

利用色彩一致性的数字伪造图像取证方法

王 波¹ 孔祥维¹ 尤新刚^{1,2}

¹大连理工大学信息安全研究中心, 大连 116024 ²北京电子技术应用研究所, 北京 100091

摘 要: 数字相机的普及和图像处理软件的应用使得数字图像正面临着被随意篡改和伪造的威胁。针对模糊润饰后的数字伪造图像, 本文提出了一种利用色彩一致性进行定位检测的数字图像取证方法。通过从图像块中提取全局色调率, 色调变化率以及异常色调率三个特征, 来量化描述模糊润饰操作对成像系统引入的图像局部色彩一致性的破坏。利用支持向量机分类器进行原始像素点和模糊像素点的分类, 并以此对模糊润饰的伪造图像区域进行定位。实验表明, 该方法能够对模糊润饰的图像进行有效检测和定位。

关键词: 色彩一致性 数字图像取证 模糊润饰 伪造图像

Digital Forensic Method Using Color Consistency

WANG Bo¹ KONG Xiangwei¹ YOU Xingang^{1,2}

¹Information Security Research Center of Dalian University of Technology, Dalian 116024

²Beijing Institute of Electronic Technology and Application, Beijing 100091)

Abstract: The widely used digital cameras and photo editing software bring us a new problem of tampered image. Focusing on the blurring operation in digital forgeries' creation, a new digital image forensic method using color consistency is proposed for localization the tampering region in the digital forgeries. By extracting three statistical features including the ratio of global hues, hue bias and abnormal hues, the color consistency introduced by imaging pipeline and destroyed by blurring operation is described. The features are then fed to the support vector machine to classify the authentic pixels and forgeries, as well as localizing the tampering regions. Experiments show the accuracy of localizing the blurring regions in the digital image forgeries.

Key Words: color consistency digital image forensics blurring image forgery

1 引言

照片的伪造和照相技术有着同样悠久的历史。但在传统的照片伪造技术中,需要造假者具备较高的暗室操作、洗相、成像等专业技能,因此伪造照片往往不易制作。随着数字时代的到来,高清晰度数码相机的普及和功能强大、操作简单的数字图像处理软件的广泛应用,使得制作一幅伪造数字图像的门槛越来越低,即便是初学者也可以通过图像处理软件制作足以假乱真的伪造数字图像。互联网传播的便利性,也让这种伪造数字图像对社会诚信体系带来的冲击和不良影响越来越大。近年来国内外发生的一系列和篡改伪造数字图像相关的风波和纠纷提醒着我们,数字图像正面临着被随意篡改和伪造的威胁。

目前国内外学者已经对数字图像取证技术开展了研究。A C Popescu^[1]和 J Fridrich^[2]针对同幅图像的复制粘贴操作,分别提出了一种利用字典排序和遍历搜索进行伪造检测的方法。为了降低遍历搜索的计算量,前者采用了主成分分析的方法,而后者则利用了量化 DCT 系数的相关性。T T Ng 和 S F Chang 提出了一种不同图像之间拼接的模型^[3],利用图像双谱特征^[4]来对拼接图像中的突变边缘进行检测。Y Q Shi 等人则利用图像的相位连续性和小波变换等高阶统计特征^[5,6]来建立自然图像模型并检测拼接伪造图像。Hilbert-Huang 变换^[7]和相机响应函数^[8,9]等特征也被用于图像伪造检测。但这些方法只适用于检测没有经过任何后处理的拼接图像。实际中,伪造者为了达到以假乱真的目的,往往对拼接后的图像进行模糊润饰等操作。

J Fridrich 利用数码相机的模式噪声^[10]鉴别图像的真伪,提出了“数字弹道”^[11]的概念。但该方法必须在获得拍摄图像数码相机的前提下才能进行

分析,这在实际情况中常常不可能实现。[12]中提出了一种利用检测光源方向对伪造图像进行分析的方法,但其检测正确率不高。I Avci 和 S Bayram 提取与图像内容无关的特征量,来对各种图像操作进行检测^[13,14],但不能对图像的伪造部分进行定位。[15]提出了一种利用异常色调率检测和定位模糊润饰伪造图像的方法。但是该方法中存在两个经验阈值,限制了该方法对不同图像的适用性。

针对模糊润饰后伪造图像的检测和伪造区域定位问题,本文采用亮度色度模型,提取色调域的三维特征来反映图像的局部色彩一致性,并利用支持向量机来对数字图像进行取证和伪造区域定位。实验表明,该算法能够有效检测和定位出模糊润饰后数字图像的伪造区域。

2 彩色图像的色彩一致性

数字图像的成像过程,往往引入图像的局部色彩一致性^[16]。图像经历的模糊润饰操作会导致数字图像原有色彩一致性遭到破坏,本节中将从多个方面对其进行分析。

在如图 1 的图像成像模型中,CFA 插值过程与 JPEG 是导致图像局部色彩一致性的重要原因。出于相机制作成本和制作工艺水准的要求,目前大部分数码相机只采用一个单色的感光器件记录光强,其前端放置一个组合的分光镜片即颜色滤波阵列(CFA)。因此每个像素点只能采集一种颜色,其他颜色需要利用颜色插值算法进行填充。这个插值过程便引入了局部的色彩一致性^[17]。而 JPEG 为了达到更高的压缩比率,往往采用对亮度和色度的不均匀采样,以及不同的量化矩阵。由于对色度的高比例采样,以及相对较大和平滑的量化步长, JPEG 压缩进一步加强了这种局部色彩一致性。

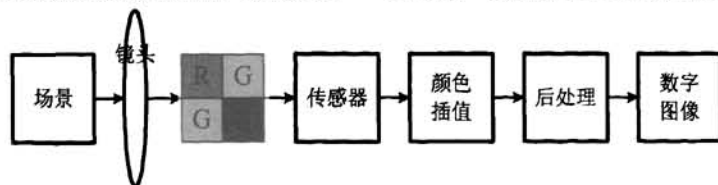


图 1 数字相机成像模型

Fig.1 Imaging Model in Digital Cameras

本文采用了 HSI 色彩模型对色彩一致性进行描

述。HSI 色彩模型根据人类视觉系统的感知特性,

将图像分为仅反映亮度信息的密度，以及反映色彩信息的色调和饱和度三个基本特征量。我们用色调 H 来描述图像的局部色彩一致性。HSI 模型和 RGB 模型的转换公式如下所示。

$$H = \arctan\left(\frac{\sqrt{3}(r-b)}{(r-g)+(r-b)}\right)$$

$$S = I - \frac{\min(r, g, b)}{r+g+b} \quad (1)$$

$$I = 0.299r + 0.587g + 0.114b$$

我们用 (R, G, B) 表示一幅未经过模糊润饰操作的数字图像的像素，而用 (R', G', B') 来表示模糊润饰操作后该像素点的三通道颜色值，其中 $R'=R+\Delta r$, $G'=G+\Delta g$, 而 $B'=B+\Delta b$ 。根据式 (1)，我们很容易得到模糊前后两个像素点的色调值，如式 (2) 和式 (3)。

$$H(R, G, B) = \arctan \frac{\sqrt{3}(R-B)}{(R-B)+(R-G)} \quad (2)$$

$$H'(R', G', B') = \arctan \frac{\sqrt{3}(R'-B')}{(R'-B')+(R'-G')} \quad (3)$$

如果要使得模糊后的伪造图像仍然保持局部的色彩一致性，那么则必须满足 $H(R, G, B) = H'(R', G', B')$ ，根据三角公式变换可得应满足下式：

$$\frac{R-G}{R-B} = \frac{R'-G'}{R'-B'} \Rightarrow \frac{\Delta r - \Delta g}{\Delta r - \Delta b} = \frac{R-G}{R-B} \quad (4)$$

由于 RGB 色彩通道与其模糊操作引起的增量 Δr 、 Δg 和 Δb 是相互独立的，式 (4) 难以满足。因此，我们认为伪造图像经历的模糊润饰操作导致了数字图像原有色彩一致性遭到破坏。通过提取特征量化描述这种色彩一致性的破坏，我们可以对模糊润饰的伪造图像进行检测。

3 特征提取和分类器

在分析数字伪造图像色彩统计特性的基础上，需要提取统计特征来量化描述这种色彩一致性以及模糊对其的破坏。本文在图像块中提取了全局色调率，色调变化率以及异常色调率三个特征，并输入到支持向量机进行分类检测和定位。

3.1 特征提取

假设一幅 $m \times n$ 的图像 I ，计算其每个像素的色

调值 H ，提取下面三个特征来衡量色彩一致性。

(1) 全局色调率：

由于相机成像过程中向数字图像引入的色彩一致性，使得图像局部存在大量相同的色调。模糊伪造操作会破坏这种色彩一致性。因此，本文定义全局色调率为色调数值总数占有像素总数的百分比，来衡量图像全局色调数值的多少，其计算公式如下：

$$f_1 = \frac{|H_I|}{mn} \quad (5)$$

其中， H_I 表示图像 I 的色调值集合， $|\cdot|$ 则表示该集合的势，即该集合中不同元素的个数。

(2) 色调变化率：

全局色调率描述了一幅图像中色调一致性的整体情况，但该特征也容易受到图像内容本身的影响。本文采用高斯低通滤波器估计模糊操作对图像色彩一致性的破坏，并计算全局色调率变化情况，作为第二个特征来对色彩一致性进行衡量。该特征描述了模糊操作导致色彩一致性的偏差程度。采用高斯低通滤波器对模糊进行估计的原因在于，高斯低通滤波是最为常用的模糊滤镜之一^[14]。色调变化率的计算公式如式 (6)。

$$f_2 = \frac{|H_{F(I)}| - |H_I|}{mn} \quad (6)$$

其中， $H_{F(I)}$ 表示利用高斯滤波器处理后的图像色调集合。

(3) 异常色调率：

[15]认为模糊操作会造成图像局部出现一些和邻域存在差异的异常色调值，并提出了异常色调率的概念。异常色调是指在图像的 8×8 邻域中单独出现，并与邻域内像素色调差的最大值不超过某一范围的色调值，其集合的势反映了一幅图像的色调值在局部邻域范围内的相关性和一致性程度。对于某一像素色调值 $h_{i,j}$ ，设 $H_{i,j}^8$ 表示以 $h_{i,j}$ 为中心的 8×8 邻域内像素色调值集合（不包括中心像素）， $\Delta \hat{h}_{i,j}$ 和 $\Delta \tilde{h}_{i,j}$ 分别表示 $h_{i,j}$ 与其 8×8 邻域内像素色调差值的最大值和最小值，即：

$$\begin{cases} \Delta \hat{h}_{i,j} = \max \{ |h_{i,j} - h_{r',j'}| \} \\ \Delta \tilde{h}_{i,j} = \min \{ |h_{i,j} - h_{r',j'}| \} \end{cases}, h_{r',j'} \in H_{i,j}^8 \quad (7)$$

定义异常色调集合 H_{abn} 如下:

① $H_{abn}^0 = \emptyset$, 即异常色调集合初始值为空集;

② 假设已得 $H_{abn}^k = \{h_{abn1}, \dots, h_{abnk}\}$, 当且仅当色调 $h_{i,j}$ 满足下式时, 令 $h_{abn(k+1)} = h_{i,j}$:

$$h_{abn(k+1)} = h_{i,j} \Leftrightarrow \begin{cases} 0 < \Delta \tilde{h}_{i,j} \leq \Delta \hat{h}_{i,j} \\ \forall h_{abnm} \in H_{abn}^k, |\Delta \tilde{h}_{i,j} - \Delta \tilde{h}_{abnm}| \neq 0, 0 \leq m \leq k \end{cases} \quad (8)$$

此时, $h_{i,j}$ 为第 $k+1$ 个异常色调值;

③ 计算所有的色调值, 得到最后的异常色调值集合 H_{abn} 。

最后, 本文采用异常色调率作为第三个特征。异常色调率的定义如式 (9) 所示。

$$f_3 = \frac{|H_{abn}|}{mn} \quad (9)$$

3.2 分类器

为了对数字图像的伪造部分进行定位, 本文算法对图像进行分块检测。图像分块的大小也是影响检测正确率的因素之一。图像块过大会导致对伪造区域的定位不准确, 而过小则会由于图像数据缺乏统计意义而导致检测准确率降低。综合考虑这两方面, 算法采用的图像子块大小为 128×128 ^[10]。

在获得了三个用于描述图像色彩一致性的统计特征后, 需要分类器判断是否是伪造区域, 本文采用支持向量机 (SVM) 对其进行分类检测。SVM 的基本思想是采用内积函数定义的非线性变换将输入空间转换到高维, 使得原本低维空间中线性不可分的问题变得高维线性可分, 然后在这个高维空间中求解广义最优分类面^[18]。本文选择支持向量机中的非线性 RBF 核函数 C-SVC (C-Support Vector Classification) 作为算法的分类器。分类器中的参数通过格形搜索的交叉校验来获得最优值, 其搜索范围为[19]中的推荐值。

4 实验与讨论

4.1 实验样本和参数

实验采用了 12 种品牌的 33 种型号相机拍摄的数码照片图像进行测试, 其中包含了市场上主流的 Canon、Nikon、Kodak、Sony 以及 Olympus 等品牌。我们从中随机挑选了 13 个相机拍摄的 156 幅照片图像作为原始图像, 而对应的模糊图像则用 Matlab 中的滤波函数生成, 以此两种图像作为训练图像。测试采用的原始图像为剩余 20 种数码相机拍摄的 864 幅照片图像, 以及 503 幅模糊后的伪造图像。这些经过模糊润饰后的伪造数字图像, 全部是在 Adobe Photoshop 8.0.1 中制作生成, 其模糊工具的参数为: 模糊半径 30-100 不等, 硬度和强度为 Photoshop 8.0.1 的默认值, 即 40% 和 50%。实验所用的支持向量机工具则来源于 C. C. Chang 等人开发的支持向量机工具包 LIBSVM^[20]。

4.2 实验结果与讨论

图 2 中第一列为相机拍摄的原始图像样本, 第二列为用该图像生成的模糊伪造数字图像, 第三列则为本文算法对模糊伪造图像的检测结果。为了衡量算法对伪造数字图像定位的正确性, 本文定义数字图像中原始像素点被检测为伪造像素的概率为虚警率, 而伪造像素点被检测为原始像素点的概率为漏检率。对所有 864 幅原始照片图像和 503 幅模糊伪造图像的统计检测结果如表 1 所示。从结果中可以看出, 算法对原始图像像素的虚警率仅为 3.7%, 而对模糊伪造图像像素的检测虚警率和漏检率则分别仅为 6.8% 和 8.9%。对模糊伪造图像的检测虚警率相对较高的原因在于, 由于伪造区域往往是非规则形状, 而本文算法的图像分块是 128×128 的矩形块。因此, 在伪造区域边缘部分, 必然存在由于分块导致的分类错误, 因此模糊伪造图像的检测虚警率高于原始图像的检测虚警率。

表 1 对所有图像像素点的统计检测结果

Tab.1 Statistical detection results of pixels in all images

	原始像素检测正确率	虚警率	伪造像素检测正确率	漏检率
原始图像	96.3%	3.7%	/	/
伪造图像	93.2%	6.8%	91.1%	8.9%

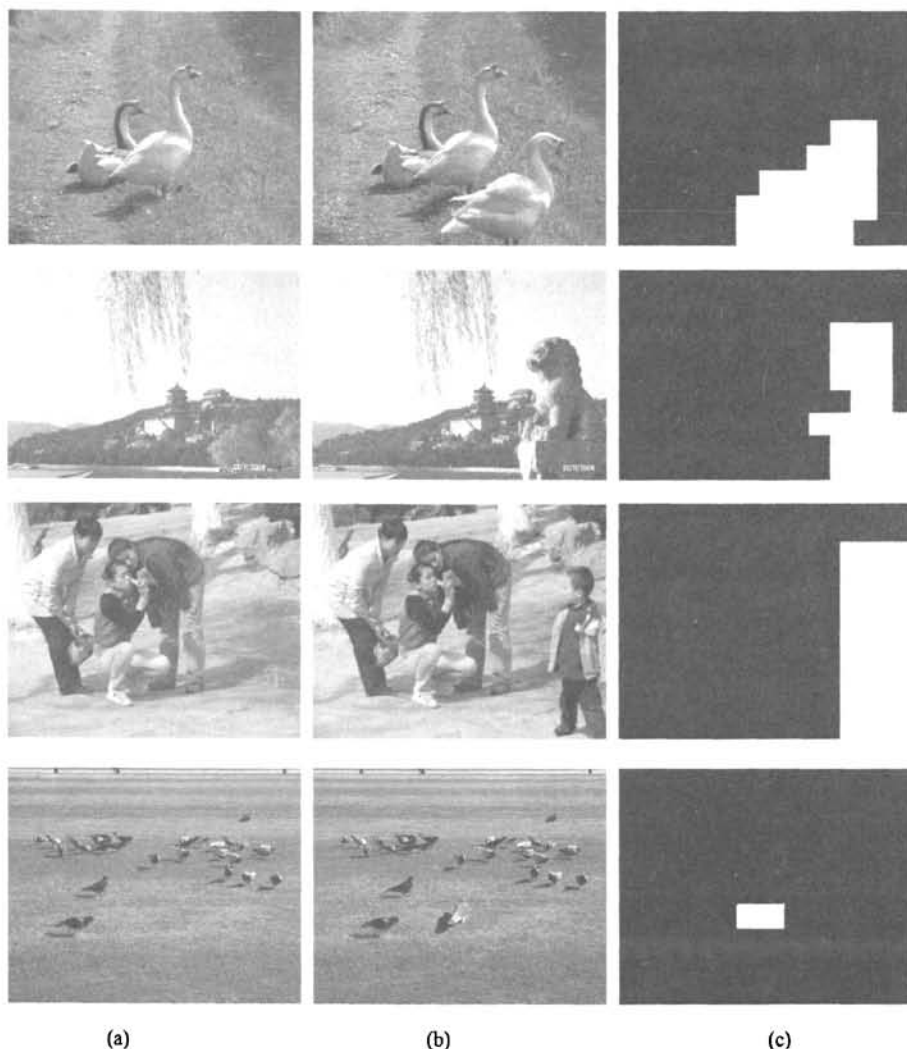


图 2 测试样本图像及检测结果。(a): 原始图像; (b): 模糊润饰后的伪造图像; (c): 对(b)的检测定位结果
Fig.2 Several samples and detection results in our experiments. (a): authentic images; (b): blurred forgeries; (c): localization results of (b)

5 结论

对模糊润饰的伪造图像进行检测定位是数字图像取证技术的重要内容之一。本文提出了一种利用色彩一致性进行数字图像取证的技术,通过全局色调率,色调变化率以及异常色调率三个特征来量化描述原始数字图像的色调一致性,以及模糊润饰操作对这种色调一致性的破坏。将图像分块后计算三个统计特征,并将其输入支持向量机进行分类,可以对模糊伪造图像进行定位检测。实验表明,算

法对模糊伪造图像的伪造区域定位具有较高的准确率,具有一定的实用性。未来的工作是进一步寻找统计特征量,对修改更小,形状不规则的伪造区域进行准确检测定位。

参考文献

[1] A. C. Popescu and H. Farid. Exposing Digital Forgeries by Detecting Duplicated Image Regions[R]. Technical Report TR 2004-515, USA: Computer Science, Dartmouth

- College, 2004.
- [2] J. Fridrich, D. Soukal, and J. Lukas. Detection of Copy-Move Forgery in Digital Images[A]. In: Proceedings of Digital Forensics Research Workshop[C], August 2003.
- [3] T. T. Ng and S. F. Chang. A Model for Image Splicing[A]. In: Proceedings of IEEE International Conference on Image Processing[C]. Singapore, Oct. 24-27, 2004: 1169~1172.
- [4] T. T. Ng and S. F. Chang. Blind Detection of Photomontage Using Higher Order Statistics[A]. In: Proceedings of IEEE International Symposium on Circuits and Systems[C], Vancouver, Canada, May 23-26, 2004: 688~691.
- [5] W. Chen, Y. Q. Shi and W. Su. Image Splicing Detection Using 2-D Phase Congruency and Statistical Moments of Characteristic Function[A]. In: Proceedings of SPIE Electronic Imaging[C], San Jose, California, USA, Jan. 28-Feb. 2, 2007: 65050R-1.
- [6] Y. Q. Shi, C. Chen, W. Chen. A Natural Image Model Approach to Splicing Detection[A]. In: Proceedings of ACM Workshop on Multimedia and Security[C], Dallas, Texas, Sep. 20-21 2007: 51~62.
- [7] D. Fu, Y. Q. Shi and W. Su. Detection of Image Splicing Based on Hilbert-Huang Transform and Moments of Characteristic Functions with Wavelet Decomposition[A]. In: Y. Q. Shi, B. Jeon (eds.): Digital Watermarking, Proceeding of 5th International Workshop on Digital Watermarking[C], Jeju Island, Korea, Nov. 8-10, 2006: 177~187.
- [8] T. T. Ng, S. F. Chang and M. P. Tsui. Using Geometry Invariants for Camera Response Function Estimation[A]. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition[C], Minneapolis, Minnesota, June 18-23, 2007: 1~8.
- [9] Y. F. Hsu and S. F. Chang. Detecting Image Splicing Using Geometry Invariants and Camera Characteristics Consistency[A]. In: Proceedings of IEEE ICME[C], Toronto, Canada, July 9-12, 2006: 549~552.
- [10] J. Lukas, J. Fridrich and M. Goljan. Detecting Digital Image Forgeries Using Sensor Pattern Noise[A]. In: Proceedings of SPIE Electronic Imaging[C], San Jose, California, USA, Jan. 15-19, 2006: 60720Y.
- [11] J. Lukas, J. Fridrich and M. Goljan. Digital 'Bullet Scratches' for Images[A]. In: Proceedings of ICIP[C], Genova, Italy, Sep. 15-17, 2005: 2637~2640.
- [12] M. K. Johnson, H. Farid. Exposing Digital Forgeries by Detecting Inconsistencies in Lighting[A]. In: Proceeding of ACM Multimedia and Security Workshop[C], New York, NY, USA, Aug. 1-2, 2005: 1~10.
- [13] I. Avcibas, S. Bayram, N. Memon, M. Ramkumar and B. Sankur. A Classifier Design for Detecting Image Manipulations[A]. In: Proceedings of IEEE ICIP[C], Singapore, Oct.24-27, 2004: 2645~2648.
- [14] S. Bayram, I. Avcibas and B. Sankur. Image Manipulation Detection[J]. Journal of Electronic Imaging, 2006, 15(4), 041102, 1~17.
- [15] 王波, 孙璐璐, 孔祥维, 尤新刚. 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.
- [16] J. Adams, K. Parulski and K. Spaulding. Color Processing in Digital Cameras[J]. IEEE Micro, 1998, 18(6): 20-30.
- [17] A. Popescu and H. Farid. Exposing Digital Forgeries in Color Filter Array Interpolated Images[J]. IEEE Transactions on Signal Processing, 2005, 53(10): 3948~3959.
- [18] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32~42.
- [19] C. W. Hsu, C. C. Chang and C. J. Lin. A Practical Guide to Support Vector Classification[DB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>, 2005.
- [20] C. C. Chang and C. J. Lin. "LIBSVM: A Library for support vector machines"[CP]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.