

基于背景噪声的图像盲篡改检测

刘丽娟^{1,2} 林小竹¹

(北京石油化工学院信息工程学院 北京 102617)¹ (北京化工大学信息科学与技术学院 北京 100026)²

摘要 数字图像在成像过程中会产生特定的背景噪声,如果两幅不同噪声的图像拼接在一起,篡改区域和其他区的噪声会有差异。提出一种基于偏度统计特性的背景噪声估计算法,其通过对图像分块计算每块的噪声标准差,从而检测出噪声异常部分以达到篡改检测的目的。算法利用DCT变换去除原图细节部分,利用偏度统计特性估计噪声,利用条件最小值法求出噪声的标准差。算法改进了迭代求条件最小值法,利用微分方法求取最小值,避免了初始值设定问题,提高了算法的准确率。实验结果表明,提出的噪声估计算法正确率高,且对拼接篡改图像篡改检测有明显效果。

关键词 图像取证,背景噪声,偏度,标准差

中图分类号 TP391 **文献标识码** A

Image Forgery Detection Using Characteristics of Background Noise

LIU Li-juan^{1,2} LIN Xiao-zhu¹

(Information Engineering College, Beijing Institute of Petrochemical Technology, Beijing 102617, China)¹

(College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China)²

Abstract There is a part of background noise in digital image which comes from imaging process. The noise characteristics between image forgery area and the other area are different if images with different noise levels are spliced together. This paper proposed a background noise estimation algorithm based on the statistical properties of skewness. We detected the forgery parts by dividing image into some sub-blocks and computing the noise variance of each ones. This algorithm removes the original image details with DCT transform, estimates noise with the statistical properties of skewness, and estimates the standard deviation of noise with condition of the minimum method. This algorithm improves iterative conditional minimum value method using differential method to calculate the minimum value. This algorithm avoids the problem of setting the initial value, and improves the accuracy of the algorithm. The experimental results show that the proposed noise estimation algorithm has high accuracy and effectiveness in detecting forgery part in spliced images.

Keywords Image forensics, Background noise, Skewness, The standard deviation

1 引言

背景噪声特性可以作为图像潜在的“水印”,通常情况下,一幅真实图像在成像过程中会引入噪声。未篡改的自然图像的噪声标准差在不同的区域大致相同,但若将两幅不同背景噪声的图像拼接在一起组成篡改图像,则篡改部分的噪声特性与原图相比会有显著的差异。

目前针对图像背景噪声的盲取证主要有4种方法。(1)利用图像去噪的方法^[1,2],即先对原始图像进行滤波来去除噪声,再用原始图像减去滤波后的图像,剩下的部分可认为是图像的背景噪声。这种方法原理和操作简单,但为了得到不含原始图像细节部分的纯噪声,对滤波器的要求比较高。(2)目前用的比较多的是小波法^[3-5],其利用的是小波变换后得到的图像高频和低频子带系数特性,认为噪声主要分布在最高频子带(HH),利用公式: $\sigma = med/0.6745$ 分块估计噪声标准差。(3)利用相机的模板噪声与待测图像匹配,该方法有一个

前提即必须已知该相机的模板噪声^[6]。(4)利用高阶统计量获得噪声标准差,Popescu等人^[7]提出把图像看成一个加性噪声模型: $Y = X + N$,一般认为噪声为高斯白噪声,峰度系数为3,利用原始和加噪图像峰度系数得到噪声标准差。该方法的前提是要知道原始图像峰度系数,但在实际检测中,我们只知道加噪图像的峰值系数。张晖等人^[8]提出一种通过邻域估计原始峰值系数方法。卢燕飞等人^[9]提出一种方法解决了原始图像峰值系数未知的难题,把待测图像当作原始图像,再在图像上人为地加上整体的高斯白噪声作为加噪图像,利用Popescu提出的方法估计噪声标准差。Zoran D^[10]和 Xunyu Pan^[11,12]等人提出一种最优值放法,该方法能够同时估计出原始峰度系数和噪声标准差,首先把图像分成 8×8 块并做DCT变换,然后对系数分块,求出峰值系数和标准差,这种方法能够有效地避免图像细节部分对噪声的干扰。

本文提出一种基于偏度统计量的噪声估计算法,该算法应用3阶矩统计特性,对图像分块计算每块的噪声标准差,从

本文受北京市属高等学校人才强教计划资助项目(PHR201107146)资助。

刘丽娟(1988-),女,硕士,主要研究方向为图像处理与模式识别,E-mail:liulijuande@163.com;林小竹(1964-),男,博士,教授,主要研究方向为图像处理与模式识别,E-mail:linxiaozhu@bipt.edu.cn。

而检测出噪声异常的部分。此算法解决了高阶统计量算法中必须已知原始信号的问题,并有效地避免了图像细节部分对噪声标准差的干扰,提高了噪声标准差估计的正确率。

2 偏度统计特性原理

对于一个随机变量 x , 它的 3 阶中心矩和偏度分别为:

$$u_3(x) = \epsilon_x \{ (x - \epsilon_x \{x\})^3 \} \quad (1)$$

$$s(x) = \frac{u_3(x)}{\sigma^3(x)} \quad (2)$$

其中,

$$\sigma^3(x) = \{\sigma^2(x)\}^{3/2} = \{\epsilon_x \{ (x - \epsilon_x \{x\})^2 \}\}^{3/2} \quad (3)$$

$$u_3 = \epsilon_x \{ (x - \epsilon_x \{x\})^3 \} \quad (4)$$

假设加噪的图像在像素域模型为 $Y = X + N$, 其中 X 为原始图像, N 为噪声, Y 为加噪后的图像, X 和 N 相互独立。那么加噪图像在频率域模型为:

$$y = x + n \quad (5)$$

其中, y, x, n 分别由 Y, X, N 频域变换得到。由于 x 和 n 相互独立, 因此 y 的方差和偏度可以写成:

$$\sigma^2(y) = \sigma^2(x) + \sigma^2(n) \quad (6)$$

$$s(y) = \frac{u_3(x) + u_3(n)}{\sigma^3(y)} = \frac{\sigma^3(x)}{\sigma^3(y)} s(x) + \frac{\sigma^3(n)}{\sigma^3(y)} s(n) \quad (7)$$

其中, $s(n) = 0$, 因为概率密度函数对称且对称轴为零的平稳噪声, 其奇次阶矩为零。由式(6)和式(7)可得:

$$s(y) = \frac{\sigma^3(x)}{\sigma^3(y)} s(x) = \left(\frac{\sigma^2(y) - \sigma^2(n)}{\sigma^2(y)} \right)^{3/2} s(x) \quad (8)$$

值得注意的是, 5 阶或者更高阶的奇次阶矩也可实现以上阶矩和噪声方差的关系, 并且更高阶矩对噪声更敏感, 但由于高阶次会消耗更多的计算时间, 因此这里只介绍偏度的计算方法。

3 背景噪声估计方法

3.1 DCT 变换

目前背景噪声估计方法大多认为噪声模型为加性高斯白噪声并与原图像相对独立, 依据统计特性可以估计图像噪声的标准差。自然图像的噪声能量主要分布在高频率子带上, 为了减小原图像中细节部分对噪声的影响, 可以先把图像从像素域变换到频率域, 只使用交流系数部分的偏度特性来估计噪声。本文参考文献[11]中方法, 选用 DCT 变换, 取交流(AC)分量各频率作为噪声估计中的 y , 图像各个块中同一频率系数组成一个子带, 每个子带的数用 y_k 表示, $k = 1, 2, \dots, 63$ 。

3.2 背景噪声估计

根据式(8)可估计原始图像的偏度和噪声标准差:

$$[\hat{s}(x), \hat{\sigma}^2(n)] = \arg \min_{s(x), \sigma^2(n)} \sum_{k=1}^K [s(y_k) - \left(\frac{\sigma^2(y_k) - \sigma^2(n)}{\sigma^2(y_k)} \right)^{3/2} s(x)]^2 \quad (9)$$

其中, 下角标 k 表示第 k 个 DCT 频率系数(DCT 变换后总共有 63 个交流系数, 这些系数按照 Z 字形顺序排序)。式(9)的最优解结果可以用 Matlab 函数 `fminsearch` 求得, 该函数可以实现无约束的非线性优化问题, 利用迭代的的思想, 选取一对初始值直到得到最优解。但选取不同的初始值可能得到的结果不同, 若对初始值设置不准确, 只能得到局部极小值, 而得不到全局最小值, 用这种方法估计标准差可能产生误差。

本章提出一种新的求解最小值方法, 利用微分求极值方

法首先对式(8)进行形式变换, 利用公式推导出方差。式(8)经过变换可得:

$$s^{2/3}(y) = s^{2/3}(x) - \frac{\sigma^2(n)}{\sigma^2(y)} \times s^{2/3}(x) \quad (10)$$

根据式(10)可令:

$$L(s^{2/3}(x), \sigma^2(n)) = \sum_{k=1}^K [s^{2/3}(y_k) - s^{2/3}(x) + \frac{\sigma^2(n)}{\sigma^2(y_k)} \times s^{2/3}(x)]^2 \quad (11)$$

当求出的 L 为最小值时, $s^{2/3}(x)$ 和 $\sigma^2(n)$ 的值为所求解。

$L(s^{2/3}(x), \sigma^2(n))$ 中两个变量的偏导数为:

$$\frac{\partial L}{\partial \sigma^2(n)} = 2 \sum_{k=1}^K [s^{2/3}(y_k) - s^{2/3}(x) + \frac{\sigma^2(n)}{\sigma^2(y_k)} \times s^{2/3}(x)] \frac{s^{2/3}(x)}{\sigma^2(y_k)} \quad (12)$$

$$\frac{\partial L}{\partial s^{2/3}(x)} = 2 \sum_{k=1}^K [s^{2/3}(y_k) - s^{2/3}(x) + \frac{\sigma^2(n)}{\sigma^2(y_k)} \times s^{2/3}(x)] \left(\frac{\sigma^2(n)}{\sigma^2(y_k)} - 1 \right) \quad (13)$$

令式(12)等于 0, 并且 $s^{2/3}(x) > 0$, 得到:

$$\sum_{k=1}^K \frac{1}{\sigma^2(y_k)} [s^{2/3}(y_k) - s^{2/3}(x) + \frac{\sigma^2(n)}{\sigma^2(y_k)} \times s^{2/3}(x)] = 0 \quad (14)$$

令式(13)等于 0 并且代入式(14)的结果, 可以得出:

$$\sum_{k=1}^K [s^{2/3}(y_k) - s^{2/3}(x) + \frac{\sigma^2(n)}{\sigma^2(y_k)} \times s^{2/3}(x)] = 0 \quad (15)$$

从式(15)可以得到:

$$\sigma^2(n) = \frac{1}{\frac{1}{K} \sum_{k=1}^K 1/\sigma^2(y_k)} - \frac{1}{s^{2/3}(x)} \frac{\sum_{k=1}^K s^{2/3}(y_k)}{\sum_{k=1}^K 1/\sigma^2(y_k)} \quad (16)$$

然后将式(16)代入式(14)中, 得到:

$$s^{2/3}(x) \left(\frac{1}{\frac{1}{K} \sum_{k=1}^K 1/\sigma^2(y_k)} - \frac{1}{s^{2/3}(x)} \frac{\sum_{k=1}^K s^{2/3}(y_k)}{\sum_{k=1}^K 1/\sigma^2(y_k)} \right) \sum_{k=1}^K \frac{1}{[\sigma^2(y_k)]^2} + \frac{\sum_{k=1}^K s^{2/3}(y_k)}{\sum_{k=1}^K \sigma^2(y_k)} + s^{2/3}(x) \sum_{k=1}^K 1/\sigma^2(y_k) = 0 \quad (17)$$

然后通过式(17)可求出 $s^{2/3}(x)$ 的值, 为了便于理解, 令 $s = s^{2/3}(y_k)$, $\sigma = \sigma^2(y_k)$, 则可以得到:

$$\hat{s}^{2/3}(x) = \frac{\langle s \rangle_K \langle 1/\sigma^2 \rangle_K - \langle s/\sigma \rangle_K \langle 1/\sigma \rangle_K}{\langle 1/\sigma^2 \rangle_K - \langle 1/\sigma \rangle_K^2} \quad (18)$$

最后把式(18)代入式(16)中, 可以得到:

$$\hat{\sigma}^2(n) = \frac{1}{\langle 1/\sigma^2 \rangle_K} - \frac{1}{\hat{s}^{2/3}(x)} \frac{\langle s \rangle_K}{\langle 1/\sigma \rangle_K} \quad (19)$$

其中, $\langle \cdot \rangle_K$ 指 K 个子带系数和的平均值。由此可求出噪声标准差估计 $\hat{\sigma}(n) = \sqrt{\hat{\sigma}^2(n)}$ 。

4 基于背景噪声的取证算法流程

首先, 介绍整体图像的噪声估计算法:

Step1 将图像分成 8×8 不重叠子块, 对每一个子块分别做 DCT 变换, 取出对应频率的交流系数, 得到 63 个交流系数子带;

Step2 把每个子带看成一个一维向量, 分别求出每个子带的偏度和方差, 得到 $s^{2/3}(y_k)$ 和 $\sigma^2(y_k)$, 其中 $k = 1, 2, \dots, 63$;

Step3 通过式(18)和式(19)估计所有子带的噪声标准差。

若要估计图像不同区域的噪声来判断篡改区域,则需要对图像做分块处理,分别计算每个块的噪声标准差。检测算法流程如下:

- Step1 将待测图像看作加噪后的图像,做分块处理;
- Step2 分别对每一块用整体算法流程估计噪声标准差;
- Step3 运用块合并算法将图像中同一噪声的块合并;
- Step4 对待测图像的噪声标准差图做后处理,运用数学形态法连接篡改区域。

算法流程中涉及到几个问题:

(1)图像分块问题。Step1 中的分块处理可以采用不重叠分块和重叠分块,为了使实验更精确,本文采用重叠分块,重叠方法为上下左右之间有一半面积重叠。设待检测图像大小为 $M \times N$, 选取每个块的大小为 $a \times b$, 那么总共能划分 $\lfloor 2M/a - 1 \rfloor \times \lfloor 2N/b - 1 \rfloor$ 个块。块的大小选取要适中,块若太大,噪声估计结果较精确,但分辨率较低,篡改区域无法精确显示;块过小则数据不充分,噪声估计结果不准确,且运算时间过长。因此应根据待测图像的实际大小来决定块大小。本文采用拼接检测库中的图片做实验,图片大小为 757×568 , 选取的块大小为 32×32 , 可以将除不尽的余下的边缘像素舍掉。

(2)块合并问题。根据每个块的噪声标准差 $\hat{\sigma}_i (i=1, \dots, n)$ 的大小,我们把图像分成几个子区域 $R_1 \cup R_2 \cup \dots \cup R_n$, 相同的子区域有相同的噪声标准差。一般认为一幅图像大部分为原图区域 R_1 , 小部分为篡改区域 R_2, \dots, R_n 。块区域合并

算法是比较块中的噪声标准差与周围邻域相似度,相似判断是基于选定的相似性阈值 T , 合并算法的步骤如下:

- Step1 给每一个块标定标签;
- Step2 按照预定顺序,检查周边区域和判断标准差的绝对值是否小于所选阈值 ($|\hat{\sigma}_i - \hat{\sigma}_j| < T$), 如果是,则把这两个块定义为同一标签;否则新建标签。
- Step3 计算每一个块与周围区域的关系,把最多一类的标签块认为是原始区域,其他都认为是篡改区域。

5 实验结果及分析

传统的小波方法将小波分解后的高频部分都认为是噪声部分,所以受图像细节的影响较大,当噪声较小时,估计值偏差较大。表 1 总结了不同标准差下所有图像的平均值,将结果与文献[4]中小波法对比。为了验证噪声估计的正确性,挑选的图像固有噪声标准差较低并且光照信息统一,然后在图像上分别施加不同标准差的高斯白噪声,选取两套未压缩格式 PNG 和 TIFF 的灰度图像库:(1)Kodak 图像库^[13]中 25 张图像;(2)UCID 图像库^[14]中 25 张图像。然后分别用本文提出的偏度法和小波估计方法估计噪声标准差,对每一个标准差下的不同图像库的 25 幅图像的估计值取平均,得到的结果如表 1 所列。从表中可以看出,本文方法优于小波估计方法,特别是在低噪声情况下,小波估计法受图像细节影响较大时,偏度法更接近实际噪声标准差。

表 1 图像库中噪声估计标准差平均值

σ		0.5	1	2	3	4	5	10	15	20	25
Kodak	小波法	1.54	1.85	2.88	3.60	4.53	5.53	10.33	15.23	20.27	25.15
	偏度法	0.98	1.51	2.64	3.23	4.21	5.16	10.17	15.17	20.16	25.13
UCID	小波法	1.56	1.91	2.87	3.55	4.54	5.48	10.29	15.24	20.24	25.11
	偏度法	0.97	1.48	2.61	3.23	4.17	5.15	10.16	15.11	20.08	25.07

为了验证本文方法估计局部噪声标准差的能力,图 1 显示了合成图像的两个例子,第一种施加条纹的高斯白噪声,每个条纹有不同的噪声标准差,值从顶部到底部变化。第二种施加棋盘格的高斯白噪声,噪声标准差统一为 20。第一幅为原始图,后两幅图是使用本章偏度法估计出的局部噪声标准差的结果。这些结果都是基于 63 个交流系数,分块大小为 32×32 的重叠滑动块。

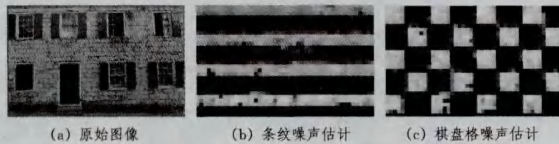


图 1 局部噪声标准差估计

本实验还随机选取了哥伦比亚大学数字图像拼接检测库^[15]中的 25 幅图像。结果证明,本文方法能够区分拼接篡改区域和非篡改区域的不同噪声特性,从而达到拼接图像篡改取证的目的。图 2 所示为实验中的一组数据。



图 2 篡改示意图

针对不同的篡改图像,要根据图像的大小来决定分块大小,块过大会导致最终标准差估计值数据不充分,难以定位篡改区域;块过小会导致块内数据不充分,估计值不准确,且运算时间过长。图 3、图 4 所示为不同分块情况下的检测结果。从图中可以看出,当选择 32×32 重叠分块时,检测结果比不重叠分块时更精确;当选择 16×16 重叠分块时,检测结果更精细,但容易受到背景的干扰,且运算时间过长;当选择 64×64 重叠分块时,检测结果更准确,但分辨率低,篡改定位误差大。因此,此幅图应选择 32×32 重叠分块策略。



图 3 不重叠与重叠分块下的检测结果



图 4 不同分块大小下的检测结果

变换的低频子带,提高了超分辨率复原图像的质量。但是以原始图像直接作为低频部分进行重建,容易忽视原始图像中低频系数的转化,使得到的重建图像出现低、高频系数不匹配的现象,从而使得到的高分辨率图像产生灰度偏移。

在对图像进行分解后,各细节侧重点是不同的(水平、垂直、对角方向的细节信息),因而对于每个部分来说,它保留的仅是与该方向相关的信息。如果此时进行双立方插值,必然会导致误差的产生;再进行小波逆变换,这样的误差便会使像素点产生重叠,导致图像模糊。根据 Haar 小波的尺度插值特性,将原图直接邻域插值后再进行 Haar 小波分解,原图可完全转化为其低频系数,且不存在任何信息损失(因为得到的高频信息均为零)。根据这个原理,可修正一部分的灰度偏差^[6]。

从表 1 的峰值信噪比 PSNR 值可以看出,经本文提出的算法处理得到的图像优于原始低分辨率图像、经双线性插值

得到的图像和经双立方插值得到的图像。

参考文献

- [1] 唐佳林,李熙莹,等.一种基于帧差法的夜间车辆检测方法[J].计算机测量与控制,2008,16(12):1811-1813
- [2] 苏秉华,金伟其,等.超分辨率图像复原及其进展[J].光学技术,2001,27(1):6-9
- [3] 陶洪久,柳健,等.基于小波变换和插值的超分辨率图像处理算法[J].武汉理工大学学报,2002,24(8):63-66
- [4] Elad M,Feuer A. Restoration of a single super resolution image from several blurred, noisy, and undersampled measured images [J]. IEEE Trans IP, 1997, 6(12):1646-1658
- [5] 戴光智,陈铁群,等.小波和插值结合应用于超声图像的超分辨率算法[J].无损检测,2009,31(4):251-255
- [6] 林虹.基于小波和插值的超分辨率图像重建算法研究[D].西安:西安电子科技大学,2005
- [7] 计算机应用研究,2009,26(3):1092-1094
- [8] Donoho D L, Johnstone J M. Ideal spatial adaptation by wavelet shrinkage[J]. Biometrika, 1994, 81(3):425-455
- [9] Mahdian B, Saic S. Using noise inconsistencies for blind image forensics[J]. Image and Vision Computing, 2009, 27(10):1497-1503
- [10] Li T, Wang M, Li T. Estimating noise parameter based on the wavelet coefficients estimation of original image[C]//2010 International Conference on Challenges in Environmental Science and Computer Engineering (CESCE). IEEE, 2010, 1:126-129
- [11] Lukáš J, Fridrich J, Goljan M. Detecting digital image forgeries using sensor pattern noise[C]//International Society for Optics and Photonics Electronic Imaging 2006. 2006, 60720Y-60720Y-11
- [12] Popescu A C, Farid H. Statistical tools for digital forensics[C]//Information Hiding. Springer Berlin Heidelberg, 2005:128-147
- [13] 张晖,张荣,尹东.使用背景噪声盲估计的图像真伪鉴别[J].中国图像图形学报,2010,15(12):1738-1741
- [14] 卢燕飞,鞠娅莉,于跃.基于图像背景噪声特性的篡改检测[J].信号处理,2012,28(9):1299-1307
- [15] Zoran D, Weiss Y. Scale invariance and noise in natural images [C]//2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009:2209-2216
- [16] Pan X, Zhang X, Lyu S. Exposing image forgery with blind noise estimation[C]//Proceedings of the Thirteenth ACM Multimedia Workshop on Multimedia and Security. ACM, 2011:15-20
- [17] Pan X, Zhang X, Lyu S. Exposing image splicing with inconsistent local noise variances[C]//2012 IEEE International Conference on Computational Photography (ICCP). IEEE, 2012:1-10
- [18] Franzen R. Kodak lossless true color image suite. 1999[OL].http://r0k.us/graphics/kodak
- [19] Schaefer G, Stich M. UCID-An Uncompressed Colour Image Database[C]//Storage and Retrieval Methods and Applications for Multimedia. SPIE, 2004:472-480
- [20] Ng T T, Chang S F. Columbia Image Splicing Detection Evaluation Dataset [EB/OL]. http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/photographers.htm

(上接第 138 页)

针对拼接检测库中的篡改图,我们还做了几组实验,实验结果如图 5 所示,左列为拼接篡改后的图像,中间列为块标准差分布图像,最右列为后处理后的图像,此图像为二值图像,白色部分为篡改区域。



图 5 拼接检测库中图像检测结果

结束语 本文提出了一种基于偏度统计特性的图像局部噪声估计算法来检测噪声异常部分即篡改区域。该算法利用 DCT 交流系数部分的偏度统计特性来估计噪声,用条件最小值法求出噪声的标准差,该算法可以有效去除原图细节的干扰,并避免了高阶统计量算法中必须已知原始信号的问题。本文改进了条件最小值算法,利用微分方法求取最小值,避免了迭代算法中初始值设定问题,提高了算法的准确率。实验结果表明,本文提出的算法针对拼接篡改图像有明显效果。

参考文献

- [1] Gou H, Swaminathan A, Wu M. Intrinsic sensor noise features for forensic analysis on scanners and scanned images[J]. IEEE Transactions on Information Forensics and Security, 2009, 4(3):476-491
- [2] 李哲,郑江滨.基于噪声分布规律的伪造图像盲检测算法[J].计