

利用 JPEG 块效应不一致性的合成图像盲检测

魏为民 唐振军

(上海大学通信与信息工程学院, 上海 200072)

摘要 合成图像的篡改对象和背景区域一般来源于不同的 JPEG 图像。为了快速有效地检测这种图像篡改,提出了一种盲检测 JPEG 合成图像的方法。首先将图像与 Laplacian 模板卷积得到二阶差分图像,沿水平(垂直)方向平均后进行离散 Fourier 变换得到归一化的频谱,基于频谱幅值构造 JPEG 块效应测度;然后将待检测图像重叠分块并计算其相应块效应测度,利用块效应不一致性检测篡改区域。实验结果表明,该方法是快速有效的。

关键词 数字取证 块效应 篡改检测 被动认证

中图法分类号: TP391.6 文献标识码: A 文章编号: 1006-8961(2009)11-2387-04

Blind Detection of Composite Images by Measuring Inconsistencies of JPEG Blocking Artifact

WEI Wei-min, TANG Zhen-jun

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

Abstract When creating a digital forgery, it is often necessary to combine several images which were originally of different JPEG compression qualities. This paper proposes a passive approach to detect digital forgeries by checking inconsistencies of blocking artifact. We calculate the second-order difference edge map by convolving the blocky image with a Laplacian kernel. Then compute the mean of the results along the horizontal (vertical) direction and obtain a vector. Discrete Fourier transform of this vector produces a normalized spectrum curve and builds a blocking artifact measure. Divide the composite image into overlapped blocks and calculate the measure of each block. The forged area can be effectively exposed by checking inconsistencies of blocking artifact. Simulation results show the efficiency of proposed algorithm.

Keywords digital forensics, blocking artifact, tampering identification, passive authentication

1 引言

数字图像很容易被篡改并广为散布。图像被动认证或盲检测的目的是在无法得到原始图像的情况下鉴别图像真伪、确定篡改位置。JPEG 压缩编码兼顾了图像存储大小和视觉质量,是最流行的图像文件格式之一,自然就成为数字图像取证的研究热点^[1-2]。

针对 JPEG 压缩特性,研究者提出了许多被动取证方法。文献[3]通过比较 8×8 不重叠分块的

块内差和块间差的直方图,判定一个非压缩的位图文件是否由 JPEG 文件转换而来,如果是,则利用最大似然估计(maximum likelihood estimation, MLE)得到相应的量化表参数。一幅 JPEG 图像被解压后再用另一个量化表重新压缩存储,这种双重 JPEG 压缩会使图像 DCT(discrete cosine transform)系数的直方图产生周期性模式,从而提供被动认证的线索^[4-6]。Farid 提出在像素域进行 JPEG 合成图像检测的方案^[7],将待检测的 JPEG 合成图像以不同的质量因子重新压缩,比较其与待检测图像的差别,用 K-S(kolmogorov-smirnov)统计工具可确定篡改区

域。需要指出的是,二次压缩并不能表明图像一定是伪造的,只能作为图像可能被篡改的一条线索。通过将块效应图像模型化为纯块信号对非块效应图像的干涉结果,文献[8]提出一种衡量图像全局块效应的方法。文献[9]利用 DCT 系数直方图的功率谱得到块效应测度,将待检测图像分为可疑区域和背景区域两部分并分别计算其块效应测度,若两者不一致则判定待检测图像可疑。

本文基于对图像二阶差分的频谱分析,定义一个新的图像块效应测量指标,用于篡改检测。将图像重叠分块计算其块效应指标,利用自然图像的邻域相似性特点来定位篡改区域。

2 基于块效应的合成篡改检测

2.1 基于频谱的块效应测度

文献[10]在检测基于双线性或者双立方插值的 JPEG 图像放大倍率的过程中,注意到由于 JPEG 块效应导致的频谱周期尖峰干扰现象,本文利用文献[10]中要抑制的干扰因素得到块效应的测量指标。设待测图像大小为 $M \times N$,若为彩色图像则逐个颜色分量计算如下:

(1)用 Laplacian 算子对图像进行卷积得到二阶差分图像 $d(i, j)$;

(2)对 $d(i, j)$ 沿水平(垂直)方向取平均得到向量 \mathbf{v} ;

(3)在 \mathbf{v} 末尾补零使其长度为 N 的最小的 2 的整数幂值,求其 DFT(discrete fourier transform),然后将频率归一化,结果记为 V 。其中在 \mathbf{v} 末尾补零是为了克服 DFT 的栅栏效应(picket-fence effect)^[11]。

(4)定义块效应测度为

$$B = \log \frac{V_{1/8} V_{2/8} V_{3/8}}{V_0} \\ = \log(V_{1/8} V_{2/8} V_{3/8}) - \log V_0 \quad (1)$$

因为 JPEG 压缩一般采用 8×8 分块,因此在图像差分后的频谱图中,1/8 倍频处(1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8)会出现峰值,图像质量因子越低,峰值越高。因为频谱关于 0.5 对称,故只取 1/8, 2/8 和 3/8 处频谱值。考虑到相同质量因子的图像块,块内像素变动越大,其二阶差分后图像的频谱直流系数越大,故式(1)中除以直流系数 V_0 ,从而可以较好地兼顾图像的视觉质量。

以图 1 所示 512×512 的灰度图像 Man 为例,用

$Q = 85$ 的质量因子将其 JPEG 压缩后的 DFT 频谱如图 2 所示,其中叉号标注的频率位置分别为 1/8, 2/8 和 3/8。因为频谱关于 0.5 对称,故只绘出其中一半的图像,另外,为便于观察将幅度较大的低频部分强行置为零。计算得到块效应测度为 6.75。



图 1 Man (512 × 512)

Fig. 1 Man (512 × 512)

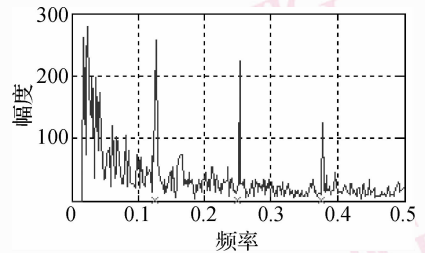


图 2 频谱曲线 (Man, $Q = 85$)

Fig. 2 Spectrum curves (Man, $Q = 85$)

在 JPEG 压缩过程中,量化表对图像质量起着决定性作用,通常与质量因子 Q 对应, Q 越小,量化表取值越大图像质量越低; Q 越大,量化表取值越小图像质量越高。以不同的 Q 压缩图 1 得到的块效应测度与质量因子的关系如图 3 所示,可见图像质量越好,块效应测度越低。用大量不同的图像测试,

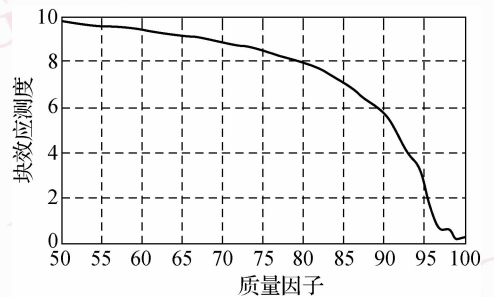


图 3 不同质量因子的块效应曲线 (Man)

Fig. 3 Block artifact measure curves (Man)

均得到图 3 类似形状的曲线。

上述方法中,块效应测度是沿图像块水平方向计算得到,亦可同时沿水平方向和垂直方向分别计算块效应测度 B_h 和 B_v ,然后将二者平均得到总的块效应测度。

2.2 基于块效应的合成篡改检测

图像合成篡改通常采用如下方法:(1)从一幅或多幅 JPEG 真实图像中提取出需要的对象;(2)选用一幅 JPEG 图像作为背景,并将提取出的对象放置在背景合适的位置上;(3)将对象的边缘进行平滑滤波,使合成图像具有更好的视觉真实性;(4)对合成图像进行 JPEG 再压缩存储。在此过程中,篡改对象和背景区域来源于不同的 JPEG 文件,其压缩质量因子有可能是不同的,然而即使是相同的质量因子,因为其来源于不同的真实图像,故其块效应测度一般是不一致的。考虑到真实图像中相邻块的像素是高度相关的,因而其块效应测度也应该是相似的,为此设计了基于块效应的合成篡改检测算法。设待测图像大小为 $M \times N$,若为彩色图像则逐个颜色分量计算如下:

(1)差分图像。用 Laplace 算子对待测图像进行卷积得到二阶差分图像 $d(i, j)$;

(2)重叠分块。将 $d(i, j)$ 从左至右、从上到下划分为互相重叠的 $b \times b$ 子块 $e(x, y)$,重叠区大小为 $l \times b$ ($l < b$),则小块的总个数为

$$T_B = \lfloor (M - l) / (b - l) \rfloor \times \lfloor (N - l) / (b - l) \rfloor \quad (2)$$

式中, $\lfloor \cdot \rfloor$ 表示向下取整,一般情况下块的尺寸应大于 16×16 。小块重叠是为了得到尽可能准确的篡改区域,重叠的区域越大,检测越精确,但耗时越多。

(3)小块检测。对每个子块 $e(x, y)$ 计算对应的块效应测度 $B(x, y)$;

(4)滤均值。将上一步骤得到的结果减去其平均值,即

$$B'(x, y) = B(x, y) - \frac{1}{T_B} \sum_x \sum_y B(x, y) \quad (3)$$

如果记 B' 的直方图为 h ,如果待测图像未经过合成,则 B' 中小于零的元素的个数与大于零的元素个数应大致相当,表现在直方图是基本左右对称,否则便可判定其经过了合成篡改。

事实上,将上述步骤(4)得到的块效应测度 B' 作图,在视觉上已经可以非常明显区分篡改区域和背景区域。当然也可采用诸如 K-S 统计类工具来自自动判别篡改区域。

3 实验结果

为检验上述算法的有效性,分别测试非压缩存储和 JPEG 压缩存储两种情况下的合成篡改检测。首先测试非压缩存储,在 Photoshop 中以如图 4(a)所示 512×512 的灰度图像 sailboat 为背景,篡改内容源于 JPEG 压缩图像,篡改后的图像如图 4(c)所示,存储为 BMP 格式。检测结果如图 4(e)所示,其中,连通在一起的较暗部分与实际篡改区域基本一致。实验中分块大小为 16×16 ,重叠区 8×16 ,硬件配置为 Pentium D CPU 2.8 GHz,内存 512 M,运算环境为 Matlab 7.1,耗时 0.5 s。全局分块效应测度为 1.73。

图 4(b)是另外一个测试例子,代表 JPEG 压缩存储,其原始质量因子为 85。为表示篡改位置的不易觉察性,从图 4(b)中取出一朵菊花,以质量因子 75 压缩后替换到原来的位置,然后以质量因子 100 存储,如图 4(d)所示,检测结果如图 4(f)所示,

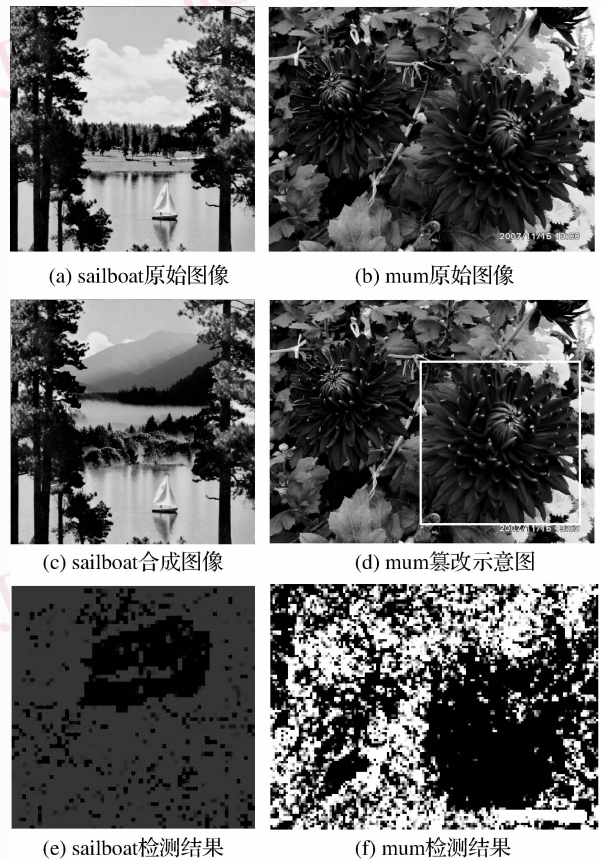


图 4 图像篡改检测

Fig. 4 Image forgery detection

基本上可明显区分篡改区域。测试的图像大小为 $1\ 200 \times 900$, 分块 16×16 , 重叠区 8×16 , 在同样的硬件和运算环境下, 耗时 5.5 s, 全局分块效应测度为 7.85。

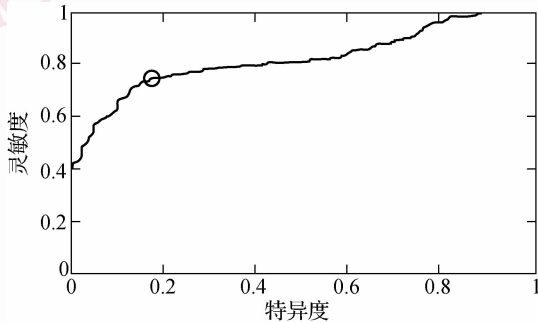


图 5 ROC 曲线 (120×120 , $Q_2 - Q_1 = 15$)

Fig. 5 ROC curves (120×120 , $Q_2 - Q_1 = 15$)

更进一步, 本文从 UCID^[12] 图像数据库中随机选择 400 幅图像, 从每一幅图像中随机取出 $b \times b$ 的图像块以质量因子 Q_1 压缩后放回原处, 然后再以质量因子 Q_2 压缩存储, Q_1 与 Q_2 为 50 到 100 之间的随机整数。以 $R = 120$, $Q_2 - Q_1 = 15$ 为例, 运用本文前述检测算法, 得到如图 5 所示的 ROC (receiver operating characteristics) 曲线, 可见该方法是比较有效的。

4 结 论

由于图像的 JPEG 压缩采用分块操作, 必然会导致块效应。本文从图像的二阶差分频谱角度, 提出一种衡量块效应的测量指标, 并将其应用到图像合成的篡改检测方面, 虽然实验表明了其有效性, 但该方法存在如下的局限性, 如待测图像中篡改的对象区域太小将很难检测, 或者存在大块的平坦区域如天空等, 二阶差分后基本为零, 亦会导致误判。正如已有的 JPEG 相关篡改检测算法^[1-2], 虽然都存在各自不尽如人意的地方, 但不能苛求一种方法解决所有的问题。特别地, 对于没有任何先验知识可资利用的盲取证, 不同的方法从不同的角度揭示篡改的痕迹, 显然增加了取证的线索。

尽管 JPEG 合成图像检测依然面临不少的难题和挑战, 但一个不争的事实是, 随着这些研究成果的

普及, 极大地提高了伪造者的造假成本, 有效地威慑了那些恶意的造假者。道高一尺, 魔高一丈, 当然也可以预见随着检测算法的不断改进, 篡改方法也会更加高明。

参考文献 (References)

- 1 Farid H. Image forgery detection [J]. IEEE Signal Processing Magazine, 2009, 26(2): 16-25.
- 2 Fridrich J. Digital image forensics [J]. IEEE Signal Processing Magazine, 2009, 26(2): 26-37.
- 3 Fan Z, Ricardo L. de Queiroz. Identification of bitmap compression history: JPEG detection and quantizer estimation [J]. IEEE Transactions on Image Processing, 2003, 12(2): 230-235.
- 4 Popescu A C. Statistical Tools for Digital Forensics [D]. Hanover, NH, USA, Dartmouth College, 2005.
- 5 Fridrich J, Lukas J. Estimation of primary quantization matrix in double compressed JPEG images [A]. In: Proceedings of Digital Forensic Research Workshop [C], Cleveland, OH, USA, 2003: 1-17.
- 6 He J, Lin Z, Wang L, et al. Detecting Doctored JPEG images Via DCT coefficient analysis [A]. In: Proceedings of 9th European Conference on Computer Vision [C], Graz, Austria: Springer, 2006: 423-435.
- 7 Farid H. Exposing digital forgeries from JPEG ghosts [J]. IEEE Transactions on Information Forensics and Security, 2009, 4(1): 154-160.
- 8 Wang Z, Bovik A C, Evans B L. Blind measurement of blocking artifacts in images [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Vancouver, Canada, 2000: 981-984.
- 9 Ye S, Sun Q, Chang Ee-chien. Detecting digital image forgeries by measuring inconsistencies of blocking artifact [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C], Beijing, China, 2007: 12-15.
- 10 Gallagher A C. Detection of linear and cubic interpolation in JPEG compressed images [A]. In: Proceedings of 2nd Canadian Conference on Computer and Robot Vision [C], Washington, DC, USA, 2005: 65-72.
- 11 Girgis A A, Ham F M. A quantitative study of pitfalls in the FFT [J]. IEEE Transactions on Aerospace and Electronic Systems, 1980, AES-16(4): 434-439.
- 12 Schaefer G, Stich M. UCID-an uncompressed colour image database [A]. In: Proceedings of SPIE, Storage and Retrieval Methods and Applications for Multimedia [C], San Jose, CA, USA, 2004: 472-480.