

# 针对 Seam-carving 图像篡改的内容自适应检测方法

吴学谦<sup>1,2</sup>, 曹 纭<sup>1,2</sup>, 赵险峰<sup>1,2</sup>, 刘长军<sup>1,2</sup>

(1. 中国科学院信息工程研究所 信息安全国家重点实验室 北京 中国 100093;

2. 中国科学院大学 网络空间安全学院 北京 中国 100093)

**摘要** 针对借助于 Seam-carving 缩放技术实现的数字图像篡改, 本文提出了一种内容自适应的检测方法。该方法考虑了 Seam-carving 的操作特点, 首先对待测图像中高度疑似篡改区域进行定位, 进而仅基于定位的区域进行特征提取及分类。该方法在特征计算时剔除了篡改发生可能性较小的区域带来的影响, 从而能够有效提高所提特征的代表性。在特征选择方面, 本文选用了扩展 Markov 特征对 Seam-carving 操作引起的像素间相关性破坏程度进行度量, 实验表明, 与先前的非自适应方法相比, 新方法在针对性检测有效性方面有着明显的优势。

**关键词** Seam-carving; 自适应检测; Markov 特征; 数字图像取证

中图分类号 TN915.08 DOI 号 10.19363/J.cnki.cn10-1380/tn.2018.11.08

## A Content Adaptive Method to Detect Seam-Carving-based Image Forgery

WU Xueqian<sup>1,2</sup>, CAO Yun<sup>1,2</sup>, ZHAO Xianfeng<sup>1,2</sup>, LIU Changjun<sup>1,2</sup>

(1. State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China;

2. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100093, China)

**Abstract** This paper proposes a content adaptive method to detect the digital image tampered by Seam-carving. Taking into account the operational characteristics of Seam-carving, in this method, the highly suspectable tampered areas in the image are located first, and then feature extraction and classification are carried out based on those located areas. The extracted features in this method are more representative for the influence of the area with less possibility is eliminated in feature calculation. Based on extended Markov feature, this paper measures the damage degree of pixels correlation caused by Seam-carving. Compared with previous non-adaptive methods, the experiment shows that the proposed method is more effective in targeted detection.

**Key words** Seam-carving; adaptive method; Markov feature; digital image forensics

### 1 前言

随着数码相机和各种图像扫描设备的普及, 数字图像逐渐成为人们生活过程中必不可少的一部分。数字图像除了记录人们的日常生活以外在司法取证, 新闻报道等方面也占据着重要的地位, 但是 Photoshop, 美图秀秀等图像篡改软件凭借着其强大多样的功能以及简单易学的操作方法迅速得到普及, 这使得数字图像的真实性越来越难以辨认, 对司法取证造成了巨大的阻碍, 使新闻报道也产生了严重

的信任危机。因此, 数字图像取证已经成为数字图像研究与应用领域里的一个重要部分。

数字图像的处理技术种类繁多, 其中图像缩放是在图像处理中很常见并且很重要的一种方法, 传统的图像缩放技术包括基于等间隔采样的缩放技术以及基于局部均值的图像缩放技术等, 这些图像缩放技术对目标图像进行非等比例缩放时, 会使图像内容发生形变, 影响图像的视觉效果。为改善缩放效果, S. Avidan 于 2007 年提出了一种基于内容感知的图像缩放技术, 即 Seam-carving 技术<sup>[1]</sup>。相较于传统的图

**通讯作者:** 刘长军, 硕士, 高级工程师, Email: liuchangjun@iie.ac.cn。

本课题得到国家自然科学基金(No.U1736214), 北京市科委项目(No.Z181100002718001), 国家重点研发计划(No.2017YFC0822704, No.2016YFB0801003, No.2016QY15Z2500), 中国科学院信息工程研究所基础前沿项目(No.Y7Z0371102)资助。

收稿日期: 2018-09-06; 修改日期: 2018-09-25; 定稿日期: 2018-09-28

像缩放技术, Seam-carving 对目标图像逐步进行能量计算, 并删除能量值最低的像素集合, 尽可能的保留用户更关心的高能量值区域。值得关注的是, 随着 Seam-carving 技术在 Adobe Photoshop CS4 上的集成应用, 这项技术的应用越来越广泛, 并且常被用于对数字图像进行恶意篡改, 例如缩小甚至删除图像中对篡改者不利的内容, 突出对篡改者有利的图像内容等。为了提高对这类篡改的识别能力, 本文提出了一种内容自适应的针对性图像篡改检测方法。

针对基于 Seam-carving 对篡改检测, 当前方法的思路还比较单一。虽然 Seam-carving 算法是比较新型的图像缩放技术, 但是其本质依然是对图像的像素进行删除/插入操作, 不可避免地会在一定程度上对图像相邻像素之间的相关性造成破坏。基于此, 目前的主流算法主要是通过检测图像像素之间的相关程度来判断图像是否经历了异常修改。在这一方向上, K. Wattanachote 提出了基于块伪影特征矩阵的特征提取方法来检测 JPEG 格式的 Seam-carving 篡改图像<sup>[2]</sup>, 该方法的核心思想也是通过提取块伪影特征矩阵中包含的像素间的相关性特征, 使用支持向量机对篡改图像和原始图像进行分类训练, 从而实现对 Seam-carving 的篡改检测, 但是该方法只能局限在 JPEG 图像的检测上, 并且其在面对比较轻微的 Seam-carving 篡改上并没有呈现出比较好的效果。A. Sarkar 提出了一种基于马尔可夫特性的特征提取方法<sup>[3]</sup>, 实验证明, 使用马尔可夫特征来描述相邻像素之间的关系具有很好的效果, 但是其只是单纯的提取了所有相邻像素之间的关联特征, 这使得其丢失了非相邻像素之间的关联特征。针对这一问题, 盛国瑞提出了一种基于扩展马尔可夫特性的特征提取方法<sup>[4]</sup>, 该方法在提取相邻像素马尔可夫特征的前提下加入了基于不同步长的非相邻像素之间的马尔可夫特性, 使得检测 Seam-carving 篡改的效果有了明显的提升; 同时, 盛国瑞还提出了一种基于 Benford 定律的特征提取方法<sup>[5]</sup>, 使用 JPEG 图像 DCT 系数的交流信号的最高位的统计特征作为检测 Seam-carving 篡改的特征, 该特征相对传统的马尔可夫特征在 Seam-carving 篡改检测正确率上也有了一个比较好的提升。此外, 杨富圣等人也提出了一种基于 LBP 特征的 Seam-carving 篡改检测方法<sup>[6]</sup>, 其利用描述图像纹理特征的局部 2 值模式算子作为检测算法的特征, 该特征相对于传统的马尔可夫特征也有比较好的效果。

经过调研发现, 目前针对 Seam-carving 的篡改检测方法都是基于整幅图像进行特征提取的, 并没

有考虑篡改区域与非篡改区域的差异性, 这导致提取的特征包含了过多的冗余信息, 影响特征的表示能力。因此, 本文提出了一种内容自适应的检测特征提取方法。该方法首先根据 Seam-carving 的操作特点, 对目标图像进行能量值计算; 然后对图像进行  $8 \times 8$  的分块, 选取其中的低能量块区域即高度疑似篡改区域进行基于马尔可夫特性的特征提取; 最后对提取特征进行支持向量机训练得到训练模型, 实现 Seam-carving 的篡改检测。该方法在很大程度上消除了所提特征中的冗余信息, 提高了针对性检测的正确率, 尤其在轻微篡改检测中表现出了明显的优势。

以下从四个部分对所提方法进行系统地论述。第 2 节对 Seam-carving 算法原理以及马尔可夫特征进行简要介绍; 第 3 节重点描述了自适应检测区域的定位及相关特征提取方法; 第 4 节给出了实验的设计过程, 并对结果进行了分析; 最后, 在第 5 节进行了总结与展望。

## 2 相关方法介绍

### 2.1 Seam-carving 图像缩放

Seam-carving 算法是 S. Avidan 于 2007 年提出的一种基于内容感知的图像缩放技术。该算法的核心在于能够实现内容感知, 对用户希望保留的关键信息进行保护, 对用户不感兴趣的次要视觉信息进行定向删除, 从而实现图像的缩放功能。该算法首先定义了一个能量矩阵来衡量图像信息的重要程度。图像中某个具体像素点能量值的函数由公式(1)给出:

$$e_1(I) = \left| \frac{\partial}{\partial x} I \right| + \left| \frac{\partial}{\partial y} I \right| \quad (1)$$

其中,  $I$  是图像强度矩阵。 $e_1(I)$  的值越小, 表明该像素点的能量越低。

Seam-carving 算法的关键点是“Seam”的确定。所谓“Seam”就是从左至右或者从上至下的由像素点组成的一条联通的路径。为了确保该路径的删除不会影响图像原本的规则矩形形状, 必须确保每条“Seam”在每一行或者每一列有且仅有一个像素点。另外, “Seam”包含的像素点应该是能量比较低的像素点, 这样可以降低在删除“Seam”的操作中, 对图像关键内容造成的破坏, 而像素点联通可以避免将一些图像关键内容中包含的低能量值的像素点包含进 Seam-carving 算法选取的“Seam”中。

Seam-carving 技术通过公式(1)计算目标图像能量分布确定图像中的重要信息, 再利用动态规划逐步找到并删除图像中具有最低能量值的最优“Seam”

来实现目标图像的缩放。下面是一个利用 Seam-carving 实现图像缩放的例子, 图 1 为 Seam-carving 缩放结果图, 其中图(a)为原始图像, 图(b)为经过 0.8 倍比例缩放的 Seam-carving 缩放图像, 图(c)为 Seam-carving 篡改操作中查找到的最优“Seam”集合。

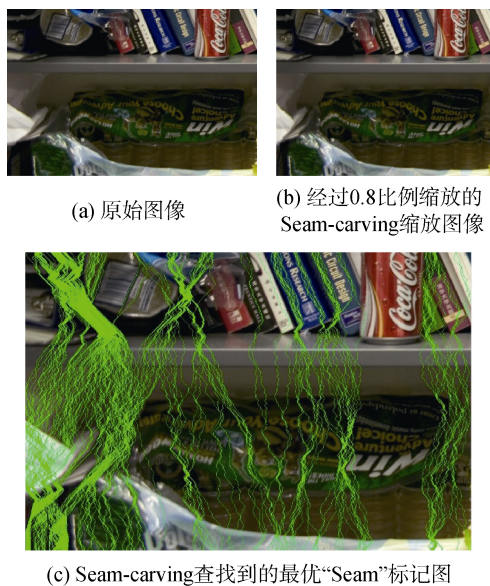


图 1 Seam-carving 缩放结果图

Figure 1 Result of Seam-carving

Seam-carving 篡改可以实现对图像的缩放, 同时利用同样的原理也可以实现图像的放大, 对于放大图像操作需要插入像素操作, Seam-carving 方法采用如下的插入算法:

(1) 将查找到的最优“Seam”删除。

(2) 在删除的每个像素的位置插入两个新的像素值。例如  $\{a_1, a_2, a_3\}$  是原图像本来的三个像素值,  $a_2$  正好是最优“Seam”中的像素值, 插入操作完成后, 相应位置的像素值变为  $a_1, b_1, b_2, a_3$ , 其中

$$b_1 = \text{round} \left\{ \frac{a_1 + a_2}{2} \right\} \quad (2)$$

$$b_2 = \text{round} \left\{ \frac{a_2 + a_3}{2} \right\} \quad (3)$$

(3) 如果选中的“Seam”恰巧在图像的边界, 那么“Seam”上的像素将被保留而只添加一个像素。比如,  $\{a_1, a_2\}$  是 2 个相邻像素,  $a_1$  或者  $a_2$  属于“Seam”并且是边界像素, 那么添加一个像素后为  $\{a_1, b, a_2\}$ , 其中

$$b = \text{round} \left\{ \frac{a_1 + a_2}{2} \right\} \quad (4)$$

## 2.2 Markov 特征

Seam-carving 操作会使得图像相邻像素之间的

相关性发生变化, 而 Markov 特征可以很好的用来表示这种相关性。图像像素之间的相关性发生变化之后同样也会影响基于  $8 \times 8$  分块的 DCT 块中的频率能量关系, 利用 Markov 特征来表征这些频率能量之间的关系也可以作为检测 Seam-carving 篡改的依据。其具体的实现方法可以分为以下几个步骤:

(1) 读取原图, 对原图进行灰度变换, 得到原图的灰度值。

(2) 对原图的灰度值进行  $8 \times 8$  的 DCT 变换, 得到相关的 DCT 系数矩阵。

(3) 将得到的 DCT 系数四舍五入进行取整, 将取整后的矩阵记为矩阵  $F$ 。

(4) 计算  $F$  水平方向和垂直方向的差值矩阵:

$$F_h(u, v) = F(u, v) - F(u+1, v) \quad (5)$$

$$F_v(u, v) = F(u, v) - F(u, v+1) \quad (6)$$

为了降低计算复杂度对差值进行进一步处理, 设定一个阈值  $T$  (正整数), 如果差值矩阵中的元素大于  $T$ , 则把其赋值为  $T$ , 若其小于  $-T$ , 则将其赋值为  $-T$ 。因为大于  $T$  值或小于  $-T$  值的值占比例比较小, 所以舍弃这些值对检测结果影响很小。经过大量实验,  $T$  值取 4 有比较好的效果。

(5) 对  $F_h$  和  $F_v$  分别在水平方向和垂直方向求转移概率矩阵:

$$P1_h = \frac{\sum_{u=1}^{su-2} \sum_{v=1}^{sv} \delta(F_h(u, v) = i, F_h(u+1, v) = j)}{\sum_{u=1}^{su-2} \sum_{v=1}^{sv} \delta(F_h(u, v) = i)} \quad (7)$$

$$P1_v = \frac{\sum_{u=1}^{su-1} \sum_{v=1}^{sv-1} \delta(F_h(u, v) = i, F_h(u, v+1) = j)}{\sum_{u=1}^{su-1} \sum_{v=1}^{sv-1} \delta(F_h(u, v) = i)} \quad (8)$$

$$P2_h = \frac{\sum_{u=1}^{su-1} \sum_{v=1}^{sv-1} \delta(F_v(u, v) = i, F_v(u+1, v) = j)}{\sum_{u=1}^{su-1} \sum_{v=1}^{sv-1} \delta(F_v(u, v) = i)} \quad (9)$$

$$P2_v = \frac{\sum_{u=1}^{su} \sum_{v=1}^{sv-2} \delta(F_v(u, v) = i, F_v(u, v+1) = j)}{\sum_{u=1}^{su} \sum_{v=1}^{sv-2} \delta(F_v(u, v) = i)} \quad (10)$$

其中,  $i$  和  $j \in \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$ ,  $S_u$  和  $S_v$  表示图像的总的行数和列数。  $\delta(x)$  是一个函数, 当括号内的条件成立时,  $\delta(x)=1$ , 如果括号内条件不成立, 则  $\delta(x)=0$ 。通过以上计算就可以得到一个  $81 \times 4 = 324$  维的特征向量。

### 3 内容自适应特征提取

#### 3.1 基于 Seam-carving 的图像篡改

图像缩放技术一直拥有非常广泛的应用场景,人们在使用数字图像的很多场合中都对目标图像的尺寸有一个限定,例如在海报制作,封面设计,网页设计中,几乎所有图像都需要进行尺寸的调整。插值算法,双线性插值算法以及双立方插值算法这些传统的图像缩放技术在缩放图像上都有很大的局限性,尤其是在非等比例的缩放中,会使得目标图像内容变形。而一般的剪切算法则很容易将目标图像的关键信息剪切破坏,尤其是目标区域比较分散的图像,根本无法使用剪切的方式来调整图像大小。因此近年来出现了很多基于图像内容的图像缩放技术,其中 Seam-carving 技术作为最典型的方法一直被人们所追捧,尤其是其在 Photoshop 中的集成,使得该技术拥有非常广泛的应用场景。

2.1 节中已经介绍过 Seam-carving 技术通过公式(1)计算目标图像能量分布确定图像中的重要信息,再利用动态规划逐步找到并删除图像中的最优“Seam”来实现目标图像的缩放。对于一幅  $N_1 \times N_2$  大小的图像,其最优“Seam”可以通过以下公式来表示:

$$s^* = \min \sum_{i=1}^{N_1} e_1(I(s_i)) \quad (11)$$

其中,  $s^*$  表示一条列的最优“Seam”,  $s$  表示最优“Seam”中的像素点集合。

最优“Seam”的具体寻找可以通过下面的方法来实现。这里以一条列“Seam”为例,从第二行开始至最后一行,对于特定的像素点(行  $i$ , 列  $j$ ),上一行与之能构成连通的只有 3 个像素点  $(i-1, j-1)$ ,  $(i-1, j)$  和  $(i-1, j+1)$ ,所以可以通过公式(5)确定选取哪一个能得到最小能量值的“Seam”。

$$M(i, j) = e_1(i, j) + \min(M(i-1, j-1), M(i-1, j), M(i-1, j+1)) \quad (12)$$

其中,  $M(i, j)$  表示经过  $(i, j)$  点的“Seam”从第一行至  $(i, j)$  处这段“Seam”的总能量值。这样,按照公式(5)从第二行开始向下计算,当计算到最后一行的各个像素点时,对应此行每个像素点的  $M(i, j)$  值也就代表了各条“Seam”最终的能量值,找出此行各个  $M(i, j)$  中的最小值并锁定位置  $(i, j)$ ,然后从  $(i, j)$  处开始回溯,逐行找到这条“Seam”在该行的元素,最终确定完整的“Seam”。找到“Seam”就可以对“Seam”进行删除操作进而实现图像的缩放。

该技术可以移除对用户不利的图像内容或者放大对用户有利的图像内容而不改变图像的其他内容,利用这一特点用户可以实现对图像的篡改。例如,可以通过把 Seam-carving 技术和插值尺度缩放结合起来的方法来放大用户指定的图像内容,首先利用传统的插值算法对目标图像进行标准的尺度放大操作,然后保持用户指定区域不变使用 Seam-carving 技术对图像进行缩放,最后得到一幅尺寸不变但是用户指定内容放大的图像。

总之, Seam-carving 的算法在图像缩放方面有着非常强大的效果,经过 Seam-carving 操作篡改过的图像往往很难通过肉眼辨别出来。另外原始图像的平滑度一般都比较自然,相邻像素之间的过渡都十分平滑,但是 Seam-carving 操作很明显会插入新的像素点或者删除已有的一些像素点,这些像素点的插入或者删除势必会以牺牲图像的平滑度为代价,使得相邻像素之间的相关性发生变化,利用这一点可以实现对 Seam-carving 篡改图像的检测,具体的检测方法会在接下来的部分进行详细的介绍。

#### 3.2 自适应区域定位

经过 Seam-carving 算法缩放过的图像都会牺牲目标图像的平滑度,由于对最优“Seam”部分进行了删除或者插入操作,所以最优“Seam”附近的像素之间的相关性发生了变化,利用其相关性的变化可以实现对 Seam-carving 的篡改检测。

图像像素之间的相关性有许多种方法可以表示,例如可以采用文献[3]中的 Markov 特征来表示,也可以采用文献[6]中的 LBP 算子来表示,还可以采用文献[7-8]中的共现矩阵来表示,这些表示方法都可以有效的表征图像相邻像素之间的关系。由于 Markov 特征是应用最广也是最被认可的特征,而扩展 Markov 特征又通过适当增加差值矩阵和转移概率矩阵步长的方法提高了 Markov 特征检测较大程度 Seam-carving 篡改图像的正确率<sup>[4]</sup>,所以本文以 Markov 特征结合扩展 Markov 特征作为训练模型的特征,以此来证明本文提出的自适应检测特征提取方法的有效性。在后续的研究中,会使用更多的特征来验证该方法的有效性。

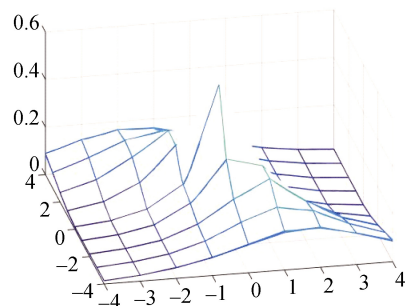
图 2 是根据公式(7)提取出的 Markov 特征图,其中图(a)是提取特征的原始图像,图(b)是提取出的 Markov 特征 3D 图。从图中可以发现,提取到的 Markov 特征分布主要集中在主对角线上,并且其在  $(0,0)$  位置达到最大值。所以本文提取了其主对角线的特征值来分析相邻像素相关性的变化。

利用公式(7)到(10)对图 2 中的原始未篡改图像以及分别经过 0.9, 0.8, 0.7, 0.6 倍 Seam-carving 篡改

过的图像进行特征提取, 然后对其主对角线的特征进行分析, 分析结果如图 3 所示。



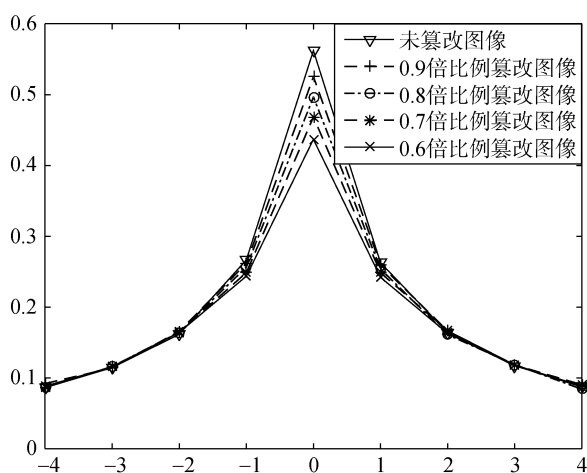
(a) 提取特征的原始图像



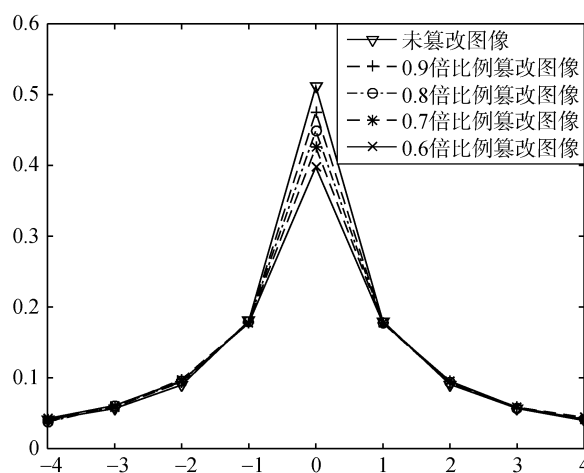
(b) 根据公式(7)提取出的Markov特征3D图

图 2 根据公式(7)提取出的 Markov 特征图

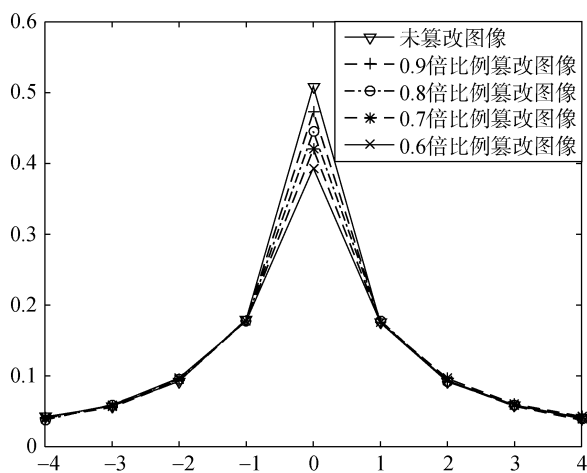
Figure 2 Markov feature map generated by formula (7)



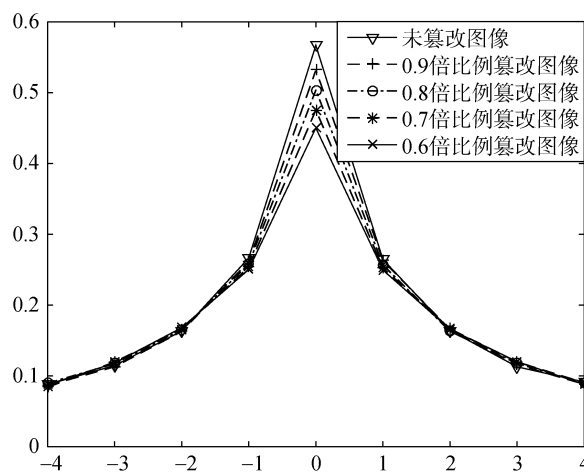
(a) 根据公式(7)提取出的Markov主对角线特征图



(b) 根据公式(8)提取出的Markov主对角线特征图



(c) 根据公式(9)提取出的Markov主对角线特征图



(d) 根据公式(10)提取出的Markov主对角线特征图

图 3 利用公式(7)到(10)提取到的 Markov 主对角线特征图

Figure 3 Markov feature map on the main diagonal generated by formula (7)-(10)

分析图 3 结果, 可以发现, 经过 Seam-carving 篡改过的图像其 Markov 主对角线的特征分布与未篡改过的原始图像相比发生了比较明显的变化, 其在

(0,0) 附近位置的所占比例随着篡改程度的增加逐步下降, 经过篡改过的图像其 Markov 特征分布会变得更加分散。本文在 2.2 节中提到, 在计算差值矩阵的



过程中设置了一个阈值  $T$ , 对小于  $-T$  值和大于  $T$  值得数值进行了舍弃, 因为这些数值所占比例较小, 对特征提取影响较小, 但是经过 Seam-carving 篡改过的图像其 Markov 特征分布会更加分散, 在  $(0,0)$  附近位置的所占比例减小, 使用阈值  $T$  筛选掉的特征值所占比例增加, 这就使得舍弃的特征值对整体特征产生了更大的影响, 从而降低了传统 Markov 特征提取方法提取到的特征的有效性。

目前关于 Seam-carving 篡改检测的方法中, 基

本都是通过提取目标图像相邻像素之间的相关性信息来训练检测模型。但是不管是基于何种方法的特征提取都是基于整幅图像的特征提取, 这些特征信息中包含了很多冗余的信息, 这些冗余信息势必会影响检测的正确率, 所以本文依据 3.1 节中的最优“Seam”提取算法, 提出了一种基于图像内容的自适应的分块算法, 来实现对图像高度疑似篡改区域的一个筛选, 尽可能的去除图像中的冗余信息, 提高检测结果的正确率。

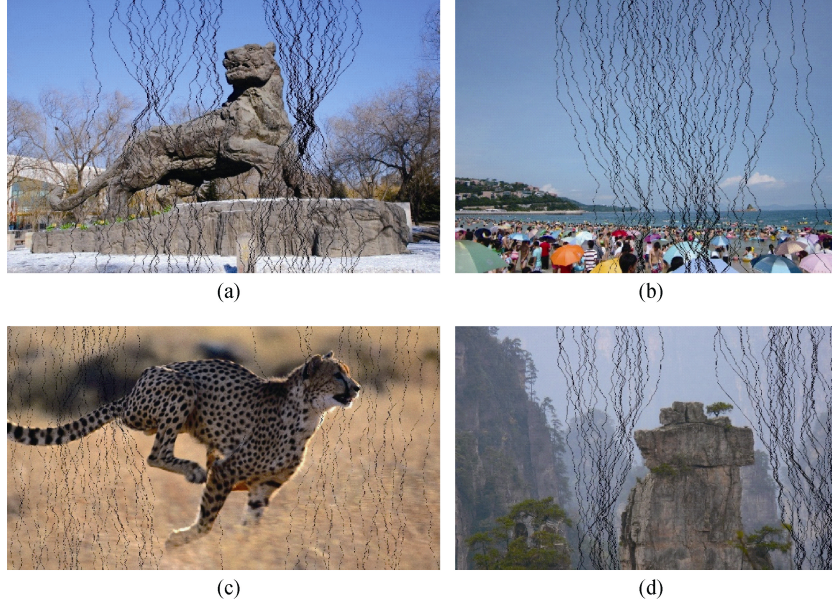


图 4 Seam-carving 算法寻找到的最优“Seam”标记图  
Figure 4 Seams created by Seam-carving algorithm

图 4 是 Seam-carving 算法找到的“Seam”, 这四张图像中的黑色线条就是 Seam-carving 算法找到的最优“Seam”, 然后在这些黑线标记的位置进行删除插入像素点的操作, 明显可以看到, 这些“Seam”比较集中, 图中很大一部分的区域是没有经过 Seam-carving 算法篡改过的。所以在提取特征之前要先对目标图像进行预处理, 首先利用公式(13)计算目标图像的能量值矩阵

$$G(x, y) = G_x(x, y) + G_y(x, y) \quad (13)$$

其中,  $G(x, y)$  是图像的梯度矩阵,  $G_x(x, y)$  和  $G_y(x, y)$  分别为图像水平梯度和垂直梯度。  $G_x(x, y)$  和  $G_y(x, y)$  可以利用以下公式计算:

$$G_x(x, y) = I(x, y) - I(x+1, y) \quad (14)$$

$$G_y(x, y) = I(x, y) - I(x, y+1) \quad (15)$$

其中,  $I(x, y)$  表示图像  $I$  在  $(x, y)$  处的像素值。然后对能量值矩阵以  $8 \times 8$  为单位进行分块处理, 得到  $M \times N$  个分块, 利用公式(16)对每一块的总体能量水

平进行一个估计。

$$E_{m,n} = \sum_{x=1}^8 \sum_{y=1}^8 G_{m,n}(x, y) \quad (16)$$

其中,  $E_{m,n}$  为第  $(m, n)$  个分块处的能量均值,  $G_{m,n}(x, y)$  为第  $(m, n)$  个分块内  $(x, y)$  处的梯度值。得到能量均值后, 对所有块的能量均值进行从小到大的排序, 排序结果如下:

$$\text{sort}E = \text{sort}(E_{m,n}) \quad (17)$$

其中,  $\text{sort}E$  为排序后的能量均值序列,  $\text{sort}$  函数为从小到大的排序函数, 由于通过 Seam-carving 算法寻找到的最优“Seam”都集中在低能量区域, 所以本文利用能量阈值  $\text{sort}E_i$  将满足公式(18)的分块标记为高度疑似篡改区域,

$$E_{m,n} < \text{sort}E_i \quad (18)$$

阈值  $\text{sort}E_i$  表示能量均值序列的第  $i$  个能量均值, 其中序号  $i$  通过公式(19)来确定:

$$i = MNP \quad (19)$$

其中,  $MN$  为能量矩阵的总分块数,  $P$  为自定义的百分比, 并且  $P \in [0.5, 0.9]$ 。之后对高度疑似篡改区域进行 Markov 特征提取, 最后将提取到的特征进行整合并用支持向量机进行分类训练得到训练模型, 从而实现对 Seam-carving 篡改的检测。

图 5 是经过自适应检测特征提取方法利用公式(7)到(10)提取的 Markov 主对角线特征图。其中  $P$  值为 0.8。相较于利用传统方法提取到的 Markov 主对角线特征比, 该方法提取的特征分布更加集中, 明显可以看到图 5 中的 Markov 主对角线特征分布在 (0,0) 附近位置具有更高的比例, 这样就可以减小利用阈值  $T$  筛选掉的特征值对提取特征所产生的影响, 提高提取特征的有效性。另外, 由于 Markov 特征的分布主要集中在 (0,0) 位置, 所以本文分析了利用自适应检测特征提取方法以及传统特征提取方法在 (0,0) 位置的特征值分布情况。

表 1 和表 2 分别是利用传统方法和自适应特征提取方法提取到的 (0,0) 位置的 Markov 特征值, 明显可以看到利用自适应特征提取方法提取到的特征分布更加集中, 其提取到的特征具有更高的有效性, 本文分别计算了不同程度的篡改图像与未篡改图像在 (0,0) 位置的 Markov 特征值的差异情况, 计算结果如表 3 和表 4 所示。

分析表 3 和表 4 的结果发现, 使用自适应特征提取方法提取的不同篡改程度图像与未篡改图像在 (0,0) 的 Markov 特征值差值比使用传统方法提取到的差值要大, 只有在使用公式(7)计算的 0.9 倍 Seam-carving 篡改与未篡改图像的差值比传统方法计算的要小。这说明了使用自适应提取方法提取的特征不仅仅分布更加集中, 在最主要特征值上, 该方法提取的篡改图像与未篡改图像的差异更大, 这使得分类效果更加有效。

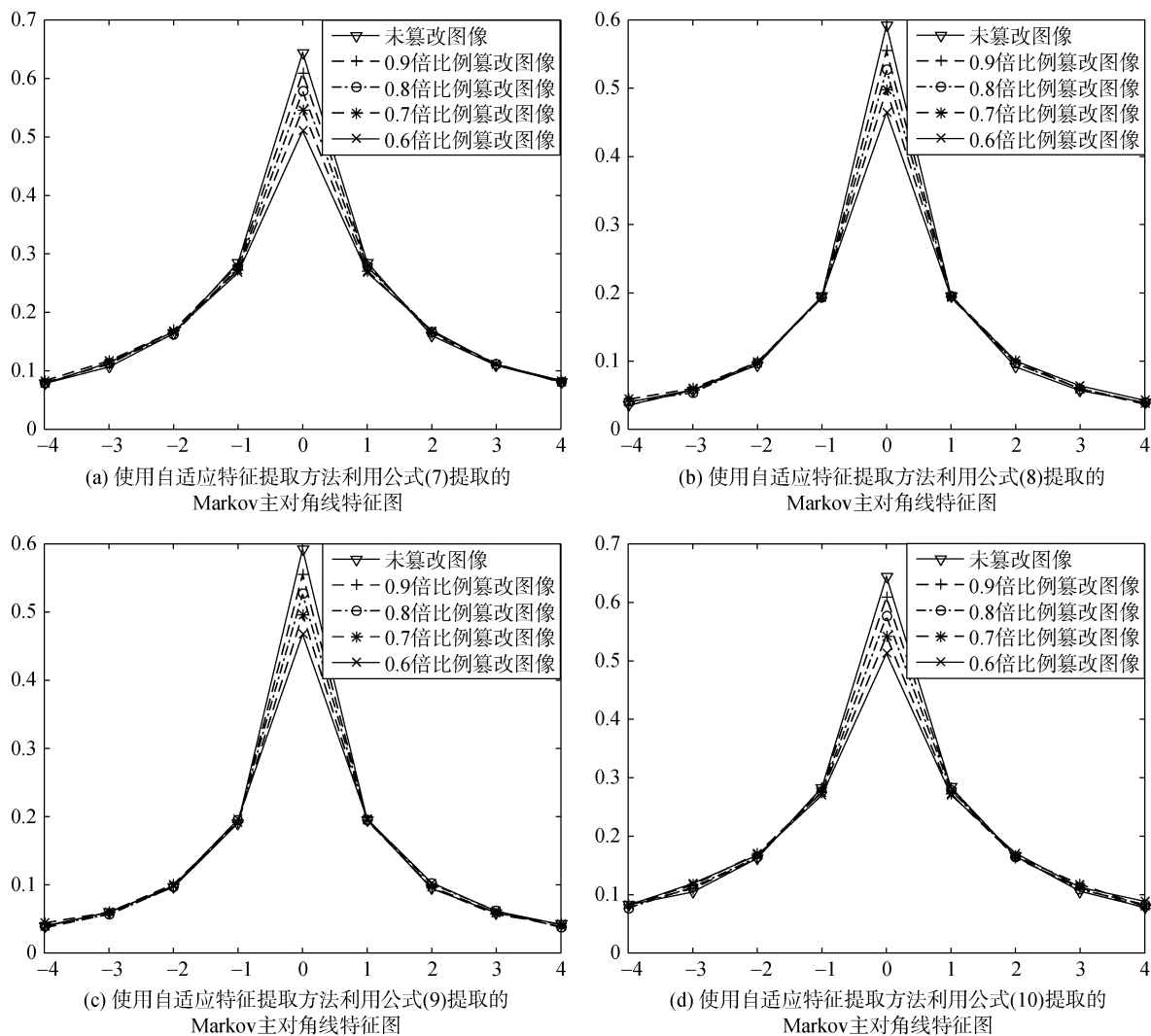


图 5 使用自适应检测特征提取方法利用公式(7)到(10)提取的 Markov 主对角线特征图

Figure 5 Markov feature map on the main diagonal generated by formula (7)-(10) using the content adaptive method

表 1 传统方法提取的(0,0)位置的 Markov 特征值  
Table 1 Markov matrix value at (0,0) using the traditional method

篡改比例 计算公式	未篡改图像	0.9	0.8	0.7	0.6
公式(7)	0.5621	0.5253	0.4967	0.4685	0.4368
公式(8)	0.5115	0.4755	0.4490	0.4254	0.3976
公式(9)	0.5090	0.4730	0.4452	0.4203	0.3935
公式(10)	0.5669	0.5320	0.5037	0.4756	0.4502

表 2 使用自适应特征提取方法提取的(0,0)位置的 Markov 特征值  
Table 2 Markov matrix value at (0,0) using the content adaptive method

篡改比例 计算公式	未篡改图像	0.9	0.8	0.7	0.6
公式(7)	0.6442	0.6096	0.5782	0.5468	0.5104
公式(8)	0.5923	0.5551	0.5266	0.4980	0.4647
公式(9)	0.5926	0.5554	0.5267	0.4967	0.4677
公式(10)	0.6443	0.6089	0.5763	0.5424	0.5125

表 3 传统方法提取的不同篡改程度图像与未篡改图像在(0,0)位置的 Markov 特征值差值表  
Table 3 The difference between Markov matrix value at (0,0) using the traditional method before and after the image is tampered

篡改比例 计算公式	0.9	0.8	0.7	0.6
公式(7)	0.0368	0.0654	0.0936	0.1253
公式(8)	0.0360	0.0625	0.0861	0.1139
公式(9)	0.0360	0.0638	0.0887	0.1155
公式(10)	0.0349	0.0632	0.0913	0.1167

表 4 自适应特征提取方法提取的不同篡改程度图像与未篡改图像在(0,0)位置的 Markov 特征值差值  
Table 4 The difference between Markov matrix value at (0,0) using the content adaptive method before and after the image is tampered

篡改比例 计算公式	0.9	0.8	0.7	0.6
公式(7)	0.0346	0.0660	0.0974	0.1338
公式(8)	0.0372	0.0657	0.0943	0.1276
公式(9)	0.0372	0.0659	0.0959	0.1249
公式(10)	0.0354	0.0680	0.1019	0.1318

综合来看, 本文提出的自适应特征提取方法提取的特征具有更高的有效性, 并且利用该特征进行分类具有更高的准确性。另外由于轻微 Seam-carving 篡改对目标图像篡改的区域较小, 所以目前对目标图像进行整幅图像的特征提取的方法都不能有效检

测轻微 Seam-carving 篡改。而该方法通过对提取特征中冗余信息的去除在轻微 Seam-carving 篡改方面具有非常明显的效果提升。

### 3.3 检测步骤

本文提出检测方法的具体步骤详细描述如下:

(1) 读取训练图像集, 利用公式(13)到(15)对图像进行梯度值计算从而得到梯度矩阵, 即 2.1 节中提到的能量值矩阵。

(2) 为了使后续计算更加简便, 结果更加明了, 对获得的能量值矩阵进行峰值归一化处理, 使其分布在 0 到 1 之间。

(3) 对计算得到的归一化能量矩阵以  $8 \times 8$  为单位进行分块操作, 得到  $M \times N$  个分块。

(4) 对于分块后的归一化能量矩阵, 利用公式(16)分别计算每一块的能量平均值, 得到均值能量矩阵。

(5) 根据公式(17)对均值能量矩阵进行排序, 本文采用对均值能量矩阵求直方图的方法对其进行排序。

(6) 利用公式(18)筛选出均值能量矩阵中能量值较低的部分, 其中能量阈值通过公式(19)来确定。图 6 所示即为  $P$  值分别取 0.5, 0.6, 0.7, 0.8, 0.9 时的低能量块筛选图。

(7) 将对于低于分界值  $sortE_i$  的  $8 \times 8$  归一化能量矩阵块定义为高度疑似篡改区域, 对该区域进行 Markov 特征提取和扩展 Markov 特征提取, 为了去除单一的低能量块的影响, 在提取特征之前进行一个判断, 判断该低能量块的周围是否有其他低能量块, 如果有则对其进行特征提取, 如果没有就对这一低能量块进行舍弃。

(8) 最后, 对提取到的特征进行整合, 并利用支持向量机对其进行分类, 创建出一个分类模型。

(9) 将测试图像按照步骤(1)到(7)进行特征提取, 将提取到的特征放入分类模型中识别, 从而实现 Seam-carving 篡改的检测。

## 4 实验与结果分析

本文实验采用的数字图像来源于被广泛使用的哥伦比亚大学彩色拼接图像库 DVMM<sup>[7]</sup>, 该图像库包含 183 张 TIF 格式的原始未篡改图像以及 183 张拼接图像, 因为要检测 Seam-carving 篡改的图像, 所以只选取了其中的 183 张原始图像作为本文的实验数据。之后, 对 183 张 TIF 格式的原始未篡改图像分别按照 0.6, 0.7, 0.8, 0.9 的比例进行 Seam-carving 篡



改操作, 从而得到 732 张篡改图像, 再加上之前的 183 张原始图像, 共计 915 张实验图像。

实验选取 100 张 TIF 格式的未篡改的图像和 100 张 TIF 格式的 Seam-carving 篡改图像作为训练数据, 其中 100 张篡改图像由 0.6, 0.7, 0.8, 0.9 倍比例进行 Seam-carving 篡改的图像中各 25 张组成。然后对训练图像分别进行基于扩展 Markov 特征的特征提取以及基于自适应特征提取方法的特征提取, 将提取到的 624 维特征分别进行支持向量机分类训练, 实验中的支持向量机采用的 RBF 核的 LIBSVM<sup>[9]</sup>。这样分别得到了基于扩展 Markov 特征的提取特征训练模型以及基于自适应特征提取方法的提取特征训练模型。

利用训练得到的两个模型, 可以对测试图像集

进行测试。测试图像集由 280 张图像组成, 其中未篡改原始图像 56 张, Seam-carving 篡改过的图像 224 张。未篡改图像为哥伦比亚图像库中不包含训练集图像的 56 张随机图像, 篡改图像为经过 Seam-carving 篡改过的不包含训练集图像的 224 张随机图像, 其中经过 0.6, 0.7, 0.8, 0.9 倍 Seam-carving 篡改的图像各 56 张。测试结果如表 5 所示。

通过上述实验结果可以发现, 采用自适应检测特征提取方法的检测结果在选取阈值为 0.8 和 0.7 时, 对未篡改图像, 经过 0.6, 0.7, 0.8, 0.9 倍比例篡改过的图像上都有了明显的效果提升, 选取阈值为 0.9 时, 对未篡改图像以及经过 0.6 倍比例篡改图像的检测结果跟基于扩展 Markov 特征的方法结果相同, 对于

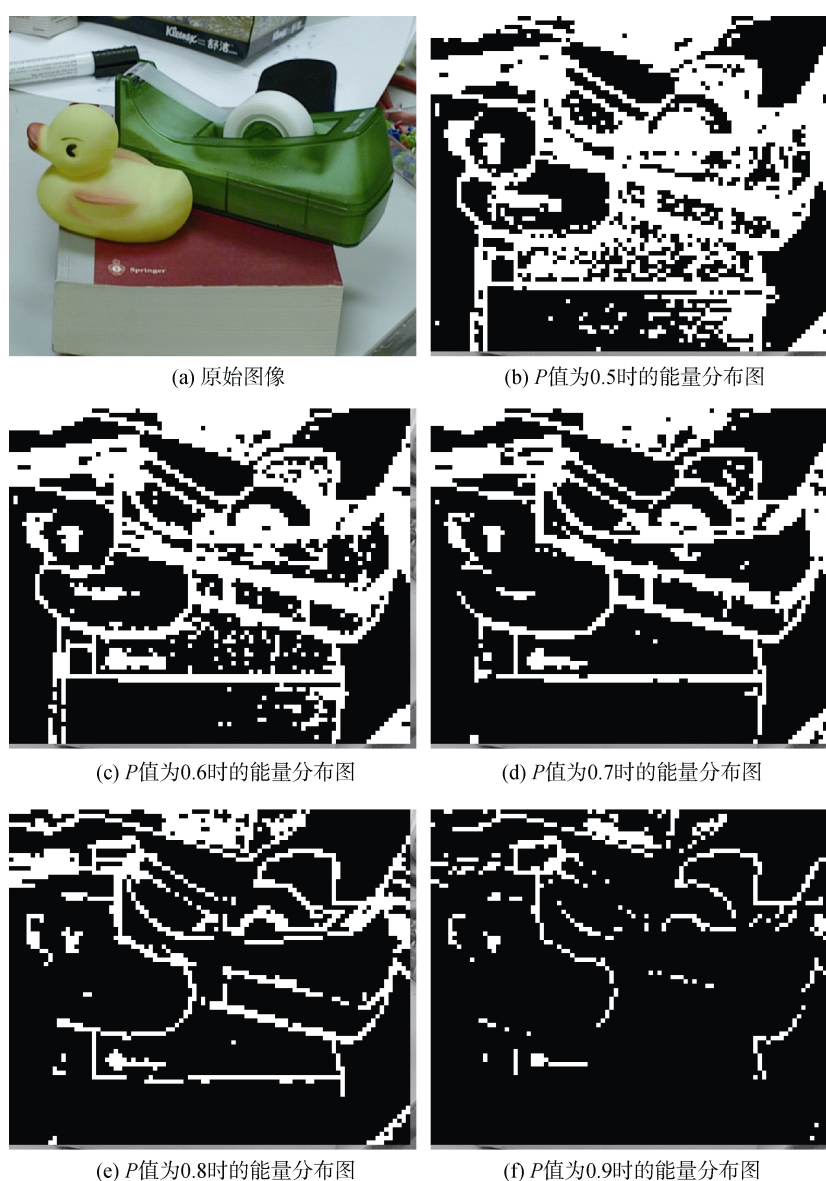


图 6  $P$  取不同值时的能量块分布图

Figure 6 Energy distribution map under different values of  $P$

(注: 图中白色部分为高于阈值的高能量块, 黑色部分为低于阈值的低能量块)

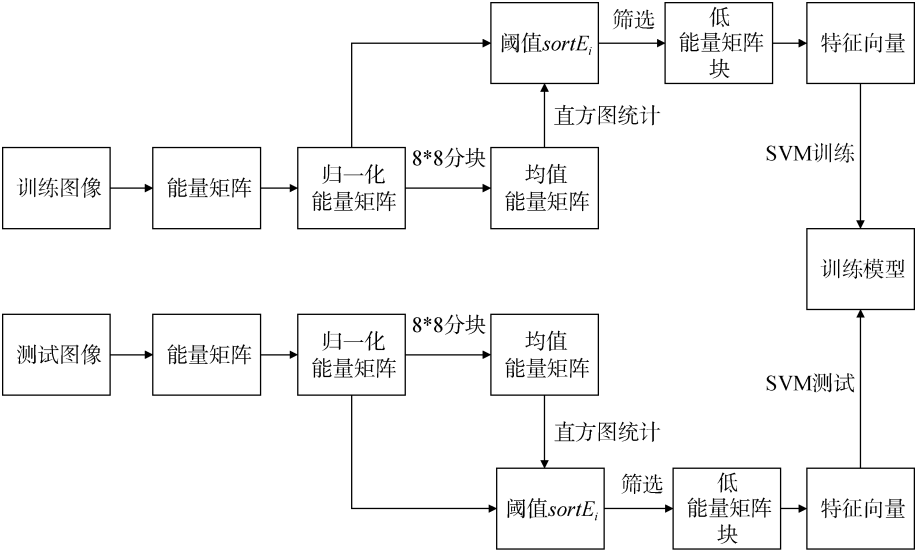


图 7 基于自适应检测分块的 Seam-carving 篡改检测流程图

Figure 7 The flow diagram of the proposed method

表 5 TIF Seam-carving 篡改图像检测正确率结果

Table 5 The detection accuracy of Seam-carving tampering images

篡改比例		未篡改图像	0.9	0.8	0.7	0.6
P 值选取						
G.R. sheng		67.85%	64.28%	85.71%	83.93%	87.50%
0.9		67.85%	78.57%	89.29%	85.71%	87.50%
0.8		69.64%	76.78%	87.50%	89.29%	91.07%
0.7		71.42%	78.57%	94.64%	85.71%	89.29%

(注: G.R. sheng 为基于扩展 Markov 特征检测结果, 其余为自适应检测特征提取方法选取不同阈值时的检测结果, 表中结果均为正确率结果)

表 6 本文算法与 G.R. sheng 算法的对比结果

Table 6 The comparison between the proposed method and G.R. sheng's method

篡改比例 算法	未篡改图像	0.9	0.8	0.7	0.6
	G.R. sheng	67.85%	64.28%	85.71%	83.93%
本文算法	69.64%	77.97%	90.489%	86.90%	89.29%

(注: 本文算法的检测结果为采用 0.7、0.8、0.9 三个阈值的平均正确率)

不同比例的篡改图像的检测结果都有所提高。另外, 分析实验结果发现三个阈值下的检测结果在对 0.9 倍比例篡改图像检测上都有比较明显的效果提升, 这是因为基于扩展 Markov 特征提取的特征在轻微 Seam-carving 篡改中包含了过多的冗余信息, 而经过自适应检测特征提取后, 其冗余信息大部分都被去除, 所以在应对轻微 Seam-carving 篡改中, 该方法具有很强的应用性。本文的方法在应对大比例 Seam-carving 篡改上的效果提升并不是非常明显, 这是因为经过大比例篡改的图像, 利用扩展 Markov 方法提取到的特征里包含的冗余信息相对较少, 而有

效信息相对较多, 经过本文方法的筛选, 在去除一部分冗余信息的过程中也可能去除了一部分有效信息, 从而使效果提升并不是十分明显, 但是通过表 6 可知, 即使是在 0.6 倍篡改比例上, 本文提出的算法相对于 G.R. sheng 提出的非自适应算法在正确率上也有也有 1.79%的提升。通过表 6 可以发现, 本文提出的自适应算法与 G.R. sheng 提出的非自适应算法相比, 在未篡改图像的检测以及经过 0.6, 0.7, 0.8, 0.9 倍比例篡改过的图像检测上都有比较明显的提升效果, 尤其是在 0.9 倍篡改比例上提高了 13.69%的正确率。

实验结果与相关分析表明, 本文提出的方法可以有效的提高基于扩展 Markov 特征的 Seam-carving 篡改检测有效性, 具体地, 在阈值为 0.8 时表现出了最好的检测效果。此外, Seam-carving 算法在过度抽取某一部分时会产生一定程度的失真现象<sup>[10]</sup>, 经过高比例 Seam-carving 篡改过的图像容易被肉眼识别出来, 所以在实际篡改过程中, 很少会选择高比例进行 Seam-carving 的篡改, 所以本文提出的算法具有很强的应用价值。

5 总结与展望

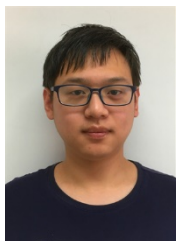
针对基于 Seam-carving 的数字图像篡改, 本文提出了一种内容自适应的检测方法。该方法在分析了 Seam-carving 操作特点的基础之上, 通过对目标图像能量矩阵进行分块处理, 筛选出高度疑似篡改区域进行特征提取, 有效提高了所提特征的有效性。本文的后续工作主要从以下两方面展开: 一方面选取/设计不同的检测特征来对本文提出方法的有效性进行进一步验证; 另一方面通过对检测图像进行预

分类, 并对不同类别图像应用不同的自适应阈值, 从而进一步提高算法的检测正确率。

**致 谢** 在此向本文成文中给予指导的老师、提供帮助的同学和给本文提出建议的评审专家表示感谢。

## 参考文献

- [1] S. Avidan, and A. Shamir, "Seam carving for content-aware image resizing," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 10, 2007.
- [2] K. Wattanachote, T.K. Shih, W.L. Chang, and H.H. Chang, "Tamper Detection of JPEG Image Due to Seam Modifications," *IEEE Transactions on Information Forensics & Security*, vol. 10, no. 12, pp. 2477-2491, 2015.
- [3] A. Sarkar, L. Nataraj, and B.S. Manjunath, "Detection of seam carving and localization of seam insertions in digital images," in *Proc. ACM Workshop on Multimedia and Security (ACM)*, pp. 107-116, 2009.
- [4] G.R. Sheng, T.G. Gao, L. Fan, L. Gao, F.S. Yang, and S. Zhang, "Seam-Carving Forgery Detection Based on Expanded Markov Features," *Journal on Communications*, vol. 35, no. 06, pp. 39-46, 2014.
- (盛国瑞, 高铁杠, 范礼, 高琳, 杨福圣, 张顺, "基于扩展马尔可夫特征的 Seam-Carving 篡改检测", *通信学报*, 2014, 35(06): 39-46。)
- [5] G.R. Sheng, and T.G. Gao, "Detection of Seam-Carving Image Based on Benford's Law for Forensic Applications," *International Journal of Digital Crime & Forensics*, vol. 8, no. 1, pp. 51-61, 2016.
- [6] F.S. Yang, and T.G. Gao, "A Method of Seam-Carving Forgery Detection Based on LBP Features," *Acta Scientiarum Naturalium Universitatis Nankaiensis*, vol. 48, no. 02, pp. 75-80, 2015.
- (杨富圣, 高铁杠。一种基于 LBP 特征的 Seam-Carving 图像篡改检测方法[J]. *南开大学学报(自然科学版)*, 2015, 48(02): 75-80。)
- [7] T.T. NG, and S.F. Chang. "A Data Set of Authentic and Spliced Image Blocks," Technical Report, 203-2004-3, Columbia University, 2004.
- [8] Y.J. Zhang, Y. Yuan, S.L. Wang, S.H. Li, and X.D. Zhao, "Detecting shifted JPEG double-compression based on conditional co-occurrence probability matrix," *Journal of Optoelectronics Laser*, vol. 23, no. 10, pp. 1932-1939, 2012.
- [9] C.C. Chang, and C.J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, pp. 27, 2011.
- [10] S. Kopf, B. Guthier, and W. Effelsberg, "Seam carving with improved edge preservation," *Proceedings of Spie the International Society for Optical Engineering*, 7542: 75420G-75420G-11, 2010.



**吴学谦** 于 2016 年在北京航空航天大学电子信息工程专业获得本科学位, 现在中国科学院大学信息工程研究所计算机技术专业攻读硕士学位, 研究兴趣包括信息对抗理论与技术、多媒体取证技术。Email: wuxueqian@iie.ac.cn



**曹纭** 于 2012 年在中国科学院软件研究所获得博士学位。现任中国科学院信息工程研究所副研究员。研究领域为多媒体内容安全。研究兴趣包括: 隐写与隐写分析、数字内容取证等。Email: caoyun@iie.ac.cn



**赵险峰** 中国科学院信息工程研究所研究员, 中国科学院大学网络空间安全学院教授, 博士生导师。2003 年于上海交通大学获博士学位, 研究方向为信息隐藏、多媒体取证与内容安全分析等。任 IJDCF、IWDW 等期刊、会议的编委、主席或委员, 任中国电子学会通信与信息安全专委会、中国图象图形学会多媒体取证与安全专委会等学术组织的委员。曾承担国家自然科学基金、国家重点研发计划、中科院战略性先导专项、部委专项等任务 40 余项, 在 IEEE TIFS、ACM IH & MMSEC 等本领域重要刊物和会议上发表论文 150 余篇, 获得与申请专利 29 项, 撰写或参与撰写著作 5 部, 主持研制的系统有重要应用, 获保密科学技术奖(部级)一等奖、中科院“朱李月华”优秀教师、ACM IH & MMSEC 最佳论文奖等荣誉。Email: zhaoxianfeng@iie.ac.cn



**刘长军** 于 2003 年获得国防科技大学硕士学位, 现在中国科学院信息工程研究所任高级工程师, 研究领域为网络体系结构与安全防护, 系统安全理论与技术。Email: liuchangjun@iie.ac.cn