

基于提升小波变换的图像篡改检测算法

柴新新, 邱晓晖

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘要: 图像在人类的感知中扮演着非常重要的角色, 所传递的信息比其他任何形式更加丰富真切, 而互联网中存在着大量的图像图片并且传播快速, 这就使得辨别图像的真伪变得尤为重要。复制粘贴操作是最为常见的图像篡改手段之一, 通过该手段可以达到修改或隐藏目标图像的目的。针对复制粘贴操作, 文中提出了一种基于提升小波变换的图像复制粘贴篡改检测算法。首先, 对图像进行提升小波变换, 提取低频分量, 并将其按像素重叠分块; 其次, 对每个小块图像进行奇异值分解得到奇异值特征向量矩阵, 并进行字典排序, 再进行欧氏距离修正; 最后, 根据偏移向量判断相似图块。实验结果表明, 该算法能够有效地减少计算量, 降低复杂度, 使检测结果更准确。

关键词: 复制粘贴; 提升小波变换; 奇异值分解; 欧氏距离

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2016)04-0078-04

doi: 10.3969/j.issn.1673-629X.2016.04.017

Image Forgery Detection Algorithm Based on Lifting Wavelet Transform

CHAI Xin-xin, QIU Xiao-hui

(College of Telecommunications & Information Engineering, Nanjing University of Posts
and Telecommunications, Nanjing 210003, China)

Abstract: Image in human perception plays a very important role, whose content is more plentiful. There are a large number of images in the Internet which spread quickly, so identifying image of authenticity is particularly important. The operation of copy-move is one of the frequently used techniques in image tampering, which can modify or hide the target image's information. For this means, a new approach toward copy-move forgery detection is proposed based on lifting wavelet transform. Firstly, the input image is decomposed into approximation (LL) subbands using lifting wavelet transform, and the LL subbands are divided into overlapping blocks. Secondly, singular value decomposition is applied to each block to obtain singular value feature matrix which is lexicographically sorted, and after sorting the matrix is fixed by Euclidean distance. Finally, the similar blocks can be judged by offset vectors. The experimental results demonstrate that the proposed method can reduce calculation amount and complexity effectively, and have more accurate detection results.

Key words: copy-move; lifting wave transform; SVD; Euclidean distance

0 引言

随着各种简易操作的图像处理软件的盛行, 无论是普通用户还是专业用户都可以轻而易举地修改图片, 并可以做到真假难辨的程度。如果图像修改只是追求更好效果仅供个人欣赏, 则无可厚非, 若是被别有用心的人利用, 后果将不堪设想。任何事物都具有两面性, 在享受信息平台带来便利的同时, 也隐藏着安全隐患。因此, 开展针对图像篡改检测技术的研究, 对于保证公众信任秩序、打击恶意修改和维护司法公正具有十分重要的意义。

图像篡改检测技术按照是否预先在数字图像中嵌

入附加信息可以分为主动检测和被动盲检测两种^[1]。

被动盲检测技术是在无任何先验条件的情况下, 进行图像的真伪鉴定^[2], 相对于主动检测技术, 不需要事先嵌入水印或签名等信息, 可以直接对图像内容的真实性进行检测, 实用性较强, 所以文中采用被动盲检测方法。

图像拼接是图像篡改常用的处理方式^[3], 复制粘贴就是最流行的方法之一, 即从一幅图像中复制一个或多个区域粘贴到同幅图像中, 以此到达隐藏目标信息的目的。针对复制粘贴操作的图像篡改被动盲检测方法主要有穷举搜索法、图像块自相关矩阵法和图像

收稿日期: 2015-07-10

修回日期: 2015-10-16

网络出版时间: 2016-03-22

基金项目: 江苏省自然科学基金 (BK2011789); 东南大学毫米波国家重点实验室开放课题 (K201318)

作者简介: 柴新新 (1988-), 女, 硕士研究生, 研究方向为智能信号处理、数字图像处理; 邱晓晖, 教授, 研究方向为现代信号中的智能信号处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160322.1520.062.html>

块匹配法^[4-5]。穷举搜索法虽然理论简单,精确度高,但运算量很大;图像块自相关矩阵法运算量较穷举搜索法小,但只能检测出较大的复制粘贴图像块,使得其应用受限;图像块匹配法虽然也存在计算量偏大的问题,但是通过其他方法可减少计算量。

1 复制粘贴篡改的图像块匹配检测

图像篡改检测块匹配算法的基本思想是通过分析待测图像中不同区域内容或像素块间的相似程度,对其进行真伪判定^[6],通常有以下五个步骤:图像预处理、图像分块、特征提取、特征匹配、篡改定位。经典的复制粘贴篡改检测流程如图1所示^[7-8]。

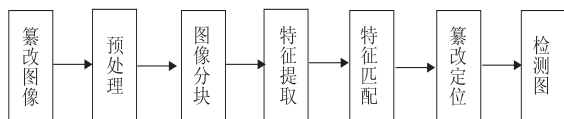


图1 经典复制粘贴检测流程

图像块匹配法为了更有效地检测到图像中的复制区域,通常在提取特征前对待检测图像进行离散余弦变换(DCT)、离散小波变换(DWT)、主成分分析(PCA)、傅里叶—梅林变换(FMT)、高斯金字塔分解等图像预处理操作,将图像空域特征转换到变换域中,对图像进行降维。整个方法的核心技术是特征提取,不同的算法在此阶段采取不同的方式提取特征。例如在空域提取图像块的均值、标准差、熵、零连通长度等统计特征,在频域提取Hu矩、Zernike矩等各种不变矩特征或SIFT、SURF、MIFT等关键点特征^[9]以及采用其他方法如奇异值分解(SVD)^[10]来提取特征等等。特征匹配阶段则根据相似度阈值、距离阈值、面积阈值等直接计算或通过聚类来完成。

不同的算法在计算量、计算复杂度以及检测结果上各有不同。文中算法采用提升小波变换和奇异值分解,在减少计算量、降低复杂度的同时可以达到更好的检测效果。

2 基于提升小波变换的图像复制粘贴篡改检测算法

小波变换具有良好的时频局部特性和多分辨率特性,因此在图像处理方面得到了广泛应用。对图像进行小波变换可以将图像划分为四个分量,即一个低频分量和三个细节分量。低频分量基本上包含了原始图像的概貌特征,细节分量分别代表了图像水平方向、垂直方向和对角方向的细节信息。通过小波变换,基于小波变换的篡改检测算法将被检图像缩减为小波变换所得的低频分量区域,缩小了检测范围,同时也保留了图像信息,但由于传统小波变换的实现是通过卷积

完成的,还存在计算复杂、运算速度慢、对内存需求量大不足^[11]。

文中为进一步降低篡改检测运算量、提高复制区域检测成功率,提出基于提升小波变换的图像复制粘贴篡改检测算法。该算法主要包括提升小波变换、奇异值分解、字典排序及修正等关键内容。

2.1 提升小波变换

作为第二代小波变换,提升小波变换既保持了传统小波的时频局部化等特性,又克服了它的局限性,其优越性表现为:

- (1) 可实现整数小波变换;
- (2) 不依赖于傅里叶变换实现小波构造;
- (3) 能在当前位置完成小波变换,节省内存;
- (4) 具有快速算法,计算更迅速。

提升小波变换的基本思想是通过一个基本小波,逐步构建出一个更具有良好性质的新小波^[12]。提升小波变换分为分解和重构两个部分,重构是分解的逆过程,如图2所示^[13]。提升算法可以分为三个步骤,即分解、预测、更新。

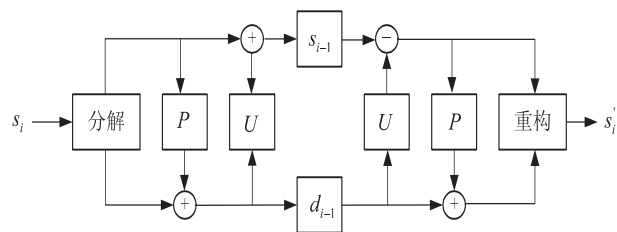


图2 提升小波变换的分解和重构示意图

(1) 分解过程。将输入信号 s_i 分为 2 个较小的子集,一般是将信号分为偶数序列 s_{i-1} 和奇数序列 d_{i-1} 两个子集,如式(1),其中 $F(s_{i-1})$ 为分解过程。

$$F(s_{i-1}) = (s_{i-1}, d_{i-1}) \quad (1)$$

(2) 预测过程。在基于原始数据相关性的基础上,用偶数序列 s_{i-1} 的预测值 $P(s_{i-1})$ 去预测奇数序列 d_{i-1} ,虽然不能准确预测 d_{i-1} ,但 $P(s_{i-1})$ 有可能很接近 d_{i-1} 。这时可以使用 $P(s_{i-1})$ 和 d_{i-1} 的差值 d'_{i-1} 来代替原来的 d_{i-1} ,如式(2),其中 P 为预测算子。

$$d'_{i-1} = d_{i-1} - P(s_{i-1}) \quad (2)$$

(3) 更新过程。经过分解步骤产生的子集的某些整体特征可能与原始数据并不一致,为了保持原始数据的这些整体特征需要一个更新过程,将更新过程用算子 U 来代替,定义为式(3)。

$$s'_{i-1} = s_{i-1} + U(d_{i-1}) \quad (3)$$

2.2 奇异值分解

奇异值分解具有三种特征,即稳定性、比例不变性和旋转不变性^[14]。对图像块进行奇异值特征向量提取,可以降低维数,减少计算量,提高抗噪声能力。奇异值分解基本理论如下:设 $A_{m \times n}$ ($m > n$) 为图像矩阵,

则存在两个正交矩阵 $U_{m \times m}$ 和 $V_{n \times n}$ 及对角矩阵 $\Sigma_{m \times n}$ 满足:

$$A = U \Sigma V^T$$

其中, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$), r 是矩阵 A 的秩。

奇异值特征向量的第一个奇异值要比其他奇异值大很多,起着主要作用,所以可以忽略一些不重要的奇异值以降低计算量。假设 Y 为奇异值特征向量矩阵:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} & 0 & \cdots & 0 \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2r} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{t1} & \sigma_{t2} & \cdots & \sigma_{tr} & 0 & \cdots & 0 \end{bmatrix} \quad (4)$$

将 Y 简化为 $Y = [\sigma_1 \sigma_2 \cdots \sigma_r]$, 则 $\eta = \frac{\sum_{i=1}^k \sigma_i}{\sum_{i=1}^r \sigma_i}$, η 为

k 维奇异值特征向量的贡献率^[3]。当 η 大于设定阈值时,每个子块的特征向量取 k 维,这样就大大降低了计算量。

$$Y' = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{t1} & \sigma_{t2} & \cdots & \sigma_{tk} \end{bmatrix} \quad (5)$$

2.3 字典排序及修正

由于字典排序过分强调了最大奇异值的作用,有可能会将奇异值相近而不是相似块排在相邻位置,从而导致误差,所以再对排序后的矩阵进行欧氏距离修正。

设 $\vec{s}_i = (s_{i1}, s_{i2}, \dots, s_{ik})$ 和 $\vec{s}_j = (s_{j1}, s_{j2}, \dots, s_{jk})$ 分别是矩阵 $S_{t \times k}$ 的两行向量,则它们之间的欧氏距离为 D_{ij} 。由于 $S_{t \times k}$ 为排序后的矩阵,只需计算某一行与其下面的 M 行之间的欧氏距离,并将与该行欧氏距离最小的一行放在该行下面,就可得到欧氏距离修正排序矩阵。

$$D_{ij} = \sqrt{\sum_{n=1}^k (s_{in} - s_{jn})^2} \quad (6)$$

2.4 算法步骤

基于提升小波变换的图像复制粘贴篡改检测算法的基本思想是先对图像进行提升小波变换提取低频分量,再将低频分量重叠分块;接着对子块进行奇异值分解,得到奇异值特征向量矩阵;然后对奇异值特征向量矩阵进行字典排序及修正;最后根据子块间的偏移向量来判断复制粘贴篡改区域。具体实现如下:

步骤 1:对待检测图像进行提升小波变换。设待

检测图像 I 为 $M \times N$ 矩阵,对图像进行 j 级提升小波变换,然后提取低频分量 LL,低频分量的矩阵大小为 $(M/2j) \times (N/2j)$ 。

步骤 2:对提取的低频分量 LL 进行重叠分块,分块大小为 $B \times B$,按像素移位,从上到下从左到右,假设分块大小 $B \times B$ 小于篡改区域,则得到的总块数 t 为 $(M/2j - B + 1) \times (N/2j - B + 1)$ 。

步骤 3:对每个子块进行奇异值分解,并变换成行向量,可得到 $(M/2j - B + 1) \times (N/2j - B + 1)$ 行 Y 矩阵。按照 2.2 节,对 Y 降维,得到新的矩阵 $Y'_{t \times k}$ 。

步骤 4:对得到的新矩阵 Y' 按字典排序得到新的矩阵 $S_{t \times k}$ 。

步骤 5:对矩阵 $S_{t \times k}$ 进行欧氏距离修正排序,得到欧氏距离修正排序矩阵 $W_{t \times k}$ 。

步骤 6:以 \vec{w}_i 表示 $W_{t \times k}$ 的第 i 行, (x_i, y_i) 表示低频分量 LL 中与 \vec{w}_i 对应的块的左上角坐标值。计算每一行与它下一行的坐标偏移向量 $\Delta x, \Delta y$ 。

$$(\Delta x, \Delta y) = (x_i - x_{i+1}, y_i - y_{i+1}) \quad (7)$$

步骤 7:对于 $\Delta x > B$ 或者 $\Delta y > B$ 的行向量进行标记,找出出现频率最多的偏移向量。

复制粘贴区域的相似块堆都具有统一的偏移向量,出现频率最多的偏移向量即可表明潜在的复制粘贴区域,对应于出现频率最多的偏移向量所在块进行标识,算法结束。

3 测试与分析

文中算法的测试环境是 MATLAB(R2011a),处理器运行频率 2.2 GHz、内存 2.0 G 的笔记本电脑。对检测结果影响大的是子块大小的设置和贡献率的选取。 B 值太小,虚警率过高; B 值太大,则会导致漏检。 B 的选取需根据情况来定,经过对多组阈值进行测试,其中以 η 大于 0.90 或 0.95, M 为 500 或者 700 效果更好。测试中的图像均为 256×256,对于彩色图像,都是将其转换成灰色图像检测的。

以参数 B 为 12, $\eta > 0.90$, M 为 500 的一组阈值进行仿真测试,结果如图 3 所示。其中图(c)、(d)是图(a)的低频分量的重构图,图中涂黑部分为复制粘贴区域。由此可见,图(d)的检测结果更为准确。

4 结束语

文中分析了图像复制粘贴篡改检测的图像块匹配法,并提出一种基于提升小波变换的篡改检测方法。与文献[4]算法比较,文中算法采用了提升小波变换,对图像降维,降低计算复杂度;在疑似区域判定的方法上做了改进,减少计算量。实验结果表明,该算法计算



图3 仿真结果图

复杂度低,检测结果更为准确。

参考文献:

[1] Ng T T, Chang S F, Sun Q. Blind detection of photomontage using higher order statistics [C]//Proceedings of the 2004 international symposium on circuits and systems. [s. l.]: IEEE, 2004.

[2] 周琳娜. 数字图像盲取证技术研究 [D]. 北京: 北京邮电大学, 2007.

[3] Ng T T, Chang S F. A model for image splicing [C]//Proc of international conference on image processing. [s. l.]: [s. n.], 2004: 24-27.

(上接第77页)

本土旅游推荐服务上的应用也验证了算法的准确性。

参考文献:

[1] 陈如明. 大数据时代的挑战, 价值与应对策略 [J]. 移动通信, 2012 (17): 14-15.

[2] 余肖生, 孙 珊. 基于网络用户信息行为的个性化推荐模型 [J]. 重庆理工大学学报: 自然科学, 2013, 27 (1): 47-50.

[3] 项 亮. 推荐系统实践 [M]. 北京: 人民邮电出版社, 2012.

[4] 刘 刚. Hadoop 应用开发技术详解 [M]. 北京: 机械工业出版社, 2014.

[5] 陶剑文. 一种分布式智能推荐系统的设计与实现 [J]. 计算机仿真, 2007, 24 (7): 296-300.

[6] 熊忠阳, 刘 芹, 张玉芳, 等. 基于项目分类的协同过滤改进算法 [J]. 计算机应用研究, 2012, 29 (2): 493-496.

[7] 黄创光, 印 鉴, 汪 静, 等. 不确定近邻的协同过滤推荐算法 [J]. 计算机学报, 2010, 33 (8): 1369-1377.

[8] Wang J, de Vries A P, Reinders M J T. Unifying user-based and item-based collaborative filtering approaches by similarity

[4] 康 丽. 数字图像内容人为篡改检测 [D]. 重庆: 西南大学, 2011.

[5] 单 薇. 基于复制粘贴的数字图像篡改检测研究 [D]. 苏州: 苏州大学, 2014.

[6] 全艳菲. 基于分块匹配的图像被动取证算法研究 [D]. 成都: 西南交通大学, 2014.

[7] Christlein V, Riess C, Angelopoulou E. On rotation invariance in copy-move forgery detection [C]//Proc of IEEE international workshop on information forensics and security. [s. l.]: IEEE, 2010: 1-6.

[8] Al-Qershi O M, Be K. Passive detection of copy-move forgery in digital images; state-of-the-art [J]. Forensic Science International, 2013, 231 (1): 284-295.

[9] Birajdar G K, Mankar V H. Digital image forgery detection using passive techniques: a survey [J]. Digital Investigation, 2013, 10 (3): 226-245.

[10] 夏 甬. 基于离散小波变换的图像篡改的检测 [D]. 北京: 中国科学院大学, 2013.

[11] 高世伟, 郭 雷, 杜亚琴, 等. 提升小波变换及其在图像处理中的应用 [J]. 计算机工程与设计, 2007, 28 (9): 2066-2069.

[12] 林樵渺. 提升格式下的小波变换在图像处理中的算法研究 [D]. 西安: 西安电子科技大学, 2005.

[13] 胡昌华, 李国华, 周 涛. 基于 MATLAB 7. x 的系统分析与设计—小波分析 [M]. 西安: 西安电子科技大学出版社, 2008.

[14] Zhang T, Wang R D. Copy-Move Forgery Detection Based on SVD in digital image [C]//Proc of 2nd international congress on image and signal processing. [s. l.]: IEEE, 2009: 1-5.

fusion [C]//Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval. New York: ACM, 2006: 501-508.

[9] Deshpande M, Karypis G. Item-based top-n recommendation algorithms [J]. ACM Transactions on Information Systems, 2004, 22 (1): 143-177.

[10] Linden G, Smith B, York J. Amazon. com recommendations: item-to-item collaborative filtering [J]. IEEE Internet Computing, 2003, 7 (1): 76-80.

[11] Liu Q. Research on some key technologies of Chinese-English machine-in translation [D]. Beijing: Peking University, 2004.

[12] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms [C]//Proceedings of the 10th international worldwide web conference. [s. l.]: [s. n.], 2001: 285-295.

[13] 刘平峰, 聂规划, 陈冬林. 基于知识的电子商务智能推荐系统平台设计 [J]. 计算机工程与应用, 2007, 43 (19): 199-201.

[14] 李 莉, 廖建伟, 欧 灵. 云计算初探 [J]. 计算机应用研究, 2010, 27 (12): 4419-4422.