

融合彩色信息与 SIFT 特征的帧内复制粘贴篡改检测^①

李小琛, 黄添强

(福建师范大学数学与信息学院, 福州 350117)
(福建省公共服务大数据挖掘与应用工程技术研究中心, 福州 350117)
通讯作者: 黄添强, E-mail: fjnu510@163.com

摘要: 近年来在同源复制粘贴篡改检测中, SIFT 特征得到了广泛的应用. 但由于该特征在提取过程中摒弃了颜色信息, 会造成一部分特征点的误匹配和漏匹配. 为此, 提出一种基于彩色信息与 SIFT 融合的 CSIFT 特征的检测方法, 在提取特征点时加入颜色不变量信息, 提高了匹配的准确性和效率. 算法首先利用结构相似度将视频帧序列分段, 提取每段序列的关键帧; 然后提取关键帧的 CSIFT 特征; 最终定位复制粘贴区域, 并利用目标跟踪算法计算篡改区域在后续帧上的位置. 通过实验验证了算法的鲁棒性, 与基于 SIFT 等特征的算法相比, 时间效率和准确性更高.

关键词: 视频篡改检测; CSIFT; 颜色不变量; 目标追踪; 复制粘贴

引用格式: 李小琛, 黄添强. 融合彩色信息与 SIFT 特征的帧内复制粘贴篡改检测. 计算机系统应用, 2018, 27(7): 11-18. <http://www.c-s-a.org.cn/1003-3254/6438.html>

Forgery Detection of Copy-Paste Video Based on Fusion of Color Information and SIFT Feature

LI Xiao-Chen, HUANG Tian-Qiang

(College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China)
(Fujian Engineering Research Center of Public Service Big Data Mining and Application, Fuzhou 350117, China)

Abstract: SIFT feature has been widely used in the homologous copy-paste forgery detection. Due to the rejection of color information, it results in some mismatching of key points, so we propose another method based on Colored Scale Invariant Feature Transform (CSIFT), a feature combining variable color information with Scale Invariant Feature Transform (SIFT) when extracting feature points decreasing the possibility of false matching and greatly reducing the time of feature point extraction and matching. In this study, we firstly segment video by Structural Similarity Index Measurement (SSIM) and extract the first frame of every part as the key frame. Then, we extract CSIFT feature points of the key frame and match feature points. After that, we locate copy-paste areas. The final step is a target tracking algorithm used to calculate the copy-paste areas in the following frame sequence. The robustness and efficiency of the algorithm is verified by experiments. Compared with algorithms based on SIFT, the proposed method has higher time efficiency and accuracy.

Key words: video forgery detection; Colored Scale Invariant Feature Transform (CSIFT); color invariant; target tracking; copy-paste

数字视频作为一种重要的多媒体数据现广泛应用于人们工作生活的各个领域, 不乏某些不法分子为

达到某些目的而对视频进行恶意的篡改, 所以如何准确高效的辨别视频真伪已成为一个亟待解决的问题^[1].

① 收稿时间: 2017-11-16; 修改时间: 2017-12-06; 采用时间: 2017-12-15; csa 在线出版时间: 2018-05-24

现在视频篡改的种类按照篡改方式可以大致分为3种: 1) 帧内篡改, 包括视频帧内对象添加与删除等篡改方式, 也称空域篡改, 本文检测的视频同源复制粘贴篡改行为是指复制一帧内的物体并粘贴到同一帧内的另一区域, 也属于帧内篡改的一种方式^[2]; 2) 帧间篡改, 也称为时域篡改, 可进一步划分为对整帧的插入、复制或删除等多种篡改方式; 3) 视频属性篡改, 如帧率转换、比特率转换等。

近年来针对视频帧内复制粘贴篡改检测, 很多检测方法被提出, 并取得一定效果, 例如文献[3]利用HOG (Histogram Of Gradient, 方向梯度直方图) 特征进行视频篡改检测, 该特征能较好的检测帧内复制粘贴篡改, 对某些时域的篡改也具有一定的检测效果, 但不能检测旋转、缩放后的粘贴区域; 文献[4,5]利用视频被篡改后需经过二次压缩这一原理, 检测二次压缩的痕迹以判断视频是否经过篡改, 这种方法在一定压缩率下是鲁棒的; 文献[6]通过篡改区域周围的纹理特征来检测篡改; 文献[7]总结了利用噪声检测视频篡改的方法; 文献[8,9]通过提取 SIFT (Scale Invariant Feature Transform, 尺度不变特征变换) 特征来检测篡改, 也取得了一定的效果。

SIFT 算法作为一种提取特征点与特征向量的经典方法, 已被广泛应用于图像、视频处理的各个领域, 尤其在图像篡改检测领域, 检测效果比较理想^[10,11]。然而在视频篡改检测领域中, 由于提取 SIFT 特征点较复杂, 造成检测的效率不高, 并且灰度化造成的信息损失使得匹配时存在一定的误匹配比率, 降低了检测准确率。因此本文利用一种彩色信息的尺度不变特征更精确地定位特征点, 在更加精确地检测篡改位置的同时, 大大降低了视频篡改检测的时间复杂度。传统的 SIFT 主要提取步骤是: 首先灰度化图片, 在尺度空间上检测极值点, 然后选择稳定的特征点并精确定位; 接着为特征点分配方向值; 最后生成特征描述子。因为在 SIFT 方法第一步中, 对图片进行了灰度化处理, 所以特征点损失了一部分的彩色信息, 增大了误匹配概率。故本文引用一种结合彩色信息与 SIFT 方法的特征点提取方法——CSIFT^[12](基于彩色不变量信息的尺度不变特征变换, a SIFT descriptor with color invariant characteristics), 将彩色信息与 SIFT 方法相结合, 更精确地提取特征点与特征向量, 减少误匹配点对的数量, 并结合结构相似度与目标追踪等方法, 提高了视频检测的效率, 优化了

检测的效果。较文献[3,8]等中的方法, 检测速率更快, 准确性更高。

1 CSIFT

现在多数特征提取的算法都是先将彩色图像转化为灰度图像后再进行配准, 色彩信息的丢失可能会引起误匹配^[13], 比如某些对象之间灰度分布相似, 但颜色分量不同, 如果仅用 SIFT 特征可能就会造成误匹配。文献[12]提出的 CSIFT 颜色不变量特征, 正好弥补了这一缺陷。

1.1 光反射模型

CSIFT 用到的颜色不变量模型是 Geusebroek 基于文献[15]和[16]中传统的 Kubelka-Munk 理论建立的, 该理论描述了物体的光谱辐射特性, 其模型称为光反射模型, 可以表达成下式:

$$E(\lambda, \vec{x}) = e(\lambda, \vec{x}) \left(1 - \rho_f(\vec{x})\right)^2 R_\infty(\lambda, \vec{x}) + e(\lambda, \vec{x}) \rho_f(\vec{x}) \quad (1)$$

其中 λ 是波长, \vec{x} 是一个二维矢量, 表示该观测点在图像中的位置, $e(\lambda, \vec{x})$ 指的是在 \vec{x} 处的光谱强度, 而 $\rho_f(\vec{x})$ 指的是该处的菲涅耳反射系数。 $R_\infty(\lambda, \vec{x})$ 表示材料反射率, $E(\lambda, \vec{x})$ 表示观测处 \vec{x} 的反射谱。一般情况下, 我们可以假设 $e(\lambda, \vec{x})$ 在各波长上保持不变而仅与观测位置 \vec{x} 有关, 于是, 我们可以将 $e(\lambda, \vec{x})$ 表示为 $i(\vec{x})$:

$$E(\lambda, \vec{x}) = i(\vec{x}) \left[\left(1 - \rho_f(\vec{x})\right)^2 R_\infty(\lambda, \vec{x}) + \rho_f(\vec{x}) \right] \quad (2)$$

1.2 CSIFT 颜色不变量特征

CSIFT 特征是基于描述物体光辐射特性的光反射模型—— $E(\lambda, \vec{x})$ 提出的, 先将 $E(\lambda, \vec{x})$ 分别对 λ 做一阶和二阶偏导, 得到一阶偏导 E_λ 和二阶偏导 $E_{\lambda\lambda}$, 如下式(3)和(4)所示:

$$E_\lambda = i(\vec{x}) \left(1 - \rho_f(\vec{x})\right)^2 \frac{\partial R_\infty(\lambda, \vec{x})}{\partial \lambda} \quad (3)$$

$$E_{\lambda\lambda} = i(\vec{x}) \left(1 - \rho_f(\vec{x})\right)^2 \frac{\partial^2 R_\infty(\lambda, \vec{x})}{\partial \lambda^2} \quad (4)$$

然后将求得的偏导 E_λ 与 $E_{\lambda\lambda}$ 相除得到颜色不变量 H 的一种表达形式。其中 $\rho_f(\vec{x}) \approx 0$, 一般可忽略, 故 H 可以表达为下式:

$$H = \left(\frac{E_\lambda}{E_{\lambda\lambda}} \right) = \frac{\partial R_\infty(\lambda, \vec{x})}{\partial \lambda} / \frac{\partial^2 R_\infty(\lambda, \vec{x})}{\partial \lambda^2} = f(R_\infty(\lambda, \vec{x})) \quad (5)$$

这里 H 是不依赖于点位置、反射面朝向、光照方向、光强大小以及菲涅耳反射系数的一种反射属性,

也是颜色不变量的一种表达形式。

根据文献[12]的推导, E 、 E_λ 和 $E_{\lambda\lambda}$ 可由RGB计算得到, 如公式(6)所示:

$$\begin{pmatrix} \hat{E} \\ \hat{E}_\lambda \\ \hat{E}_{\lambda\lambda} \end{pmatrix} = \begin{pmatrix} 0.06 & 0.63 & 0.27 \\ 0.3 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (6)$$

通过以上公式计算得到颜色不变量 H 后, 将 H 代替原始SIFT算法中的灰度图像函数 $I(x, y)$, 进行尺度空间极值检测。值得注意的是, 原始SIFT算法中根据 σ 的值计算出高斯模板矩阵的大小, 这里 σ 值越大则卷积的模板就越大, 模板越大与原图像做卷积时耗费的时间就越多, 故本文实验部分设置的 σ 值与文献[17]中设置的SIFT算法中的 σ 值保持一致, 以保证颜色不变量与SIFT的效率对比更客观, 不受 σ 参数的影响。

CSIFT是基于彩色图像的颜色不变量信息的尺度不变特征变换的一种图像特征点提取和匹配方法。与SIFT相比, 改进的CSIFT方法选用颜色不变量代替图像的灰度值进行计算, 算法复杂度并没有很明显的增加。灰度值的计算公式如下所示:

$$Gray = 0.3R + 0.59G + 0.11B \quad (7)$$

而颜色不变量的计算方法如公式(5)所示, 通过比较可见, 增加的计算量是线性的, 几乎可以忽略不计。但基于彩色图像的颜色不变量信息的尺度不变特征变换却减少了由于特征灰度化造成的彩色信息丢失而引起的特征点误匹配现象。

2 基于CSIFT的视频复制粘贴篡改检测

本文的检测算法利用上述颜色不变量信息的尺度不变特征, 同时结合K-means聚类方法和目标追踪的方法检测复制粘贴篡改行为, 算法主要步骤如图1算法框图和图2算法具体流程图所示, 首先通过比较相邻和相近帧之间的结构相似度, 将视频分段, 相似度高的相邻帧分为同一段, 然后提取每段的第一帧作为关键帧; 接着对关键帧提取该帧CSIFT特征点, 并匹配特征点; 然后通过匹配的特征点对定位复制粘贴区域; 最后利用目标跟踪算法在后续帧中追踪篡改区域, 提高算法效率。

2.1 提取关键帧

算法第一步, 将待检测视频转换为帧序列, 比较相邻和相近帧之间的结构相似度, 相似度高且连续的帧序列作为一个视频分段, 提取每一分段的第一帧作为

关键帧, 对关键帧进行进一步的检测。

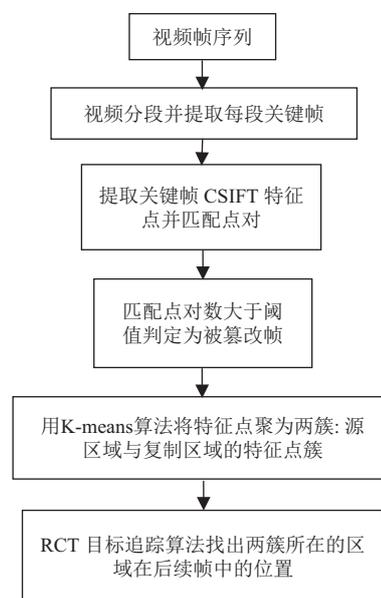


图1 本文算法框架

在这一步骤中使用的分段策略是比较帧间的结构相似度 (Structural SIMilarity, SSIM)^[18], 这是一种经典的衡量两幅图像相似度的指标, 于2004年由德州大学奥斯丁分校的图像和视频工程实验室提出, 提取SSIM的计算图像 x 和图像 y 的SSIM系数, 主要分为三个部分: 亮度比较函数 $l(x, y)$ 的计算, 对比度比较函数 $c(x, y)$ 的计算, 以及结构比较函数 $s(x, y)$ 的计算。SSIM计算公式由以上三个部分合并组成, 如下式:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (8)$$

其中,

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (9)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (10)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (11)$$

μ_x 、 σ_x 分别表示图像 x 的均值和方差, μ_y 、 σ_y 分别表示图像 y 的均值和方差, σ_{xy} 则表示两幅图像 x 与 y 的协方差。 C_1 、 C_2 、 C_3 都取非常小的数值, 目的是为了防止发生分母为0的情况。 α 、 β 、 γ 是用来调整三个分量 $l(x, y)$ 、 $c(x, y)$ 和 $s(x, y)$ 的比重。为使计算更简单, Wang等人^[18]将 α 、 β 、 γ 都设为1, 另外设置 $C_3=C_2/2$, 于是得

到一个特殊的 SSIM 系数计算公式:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (12)$$

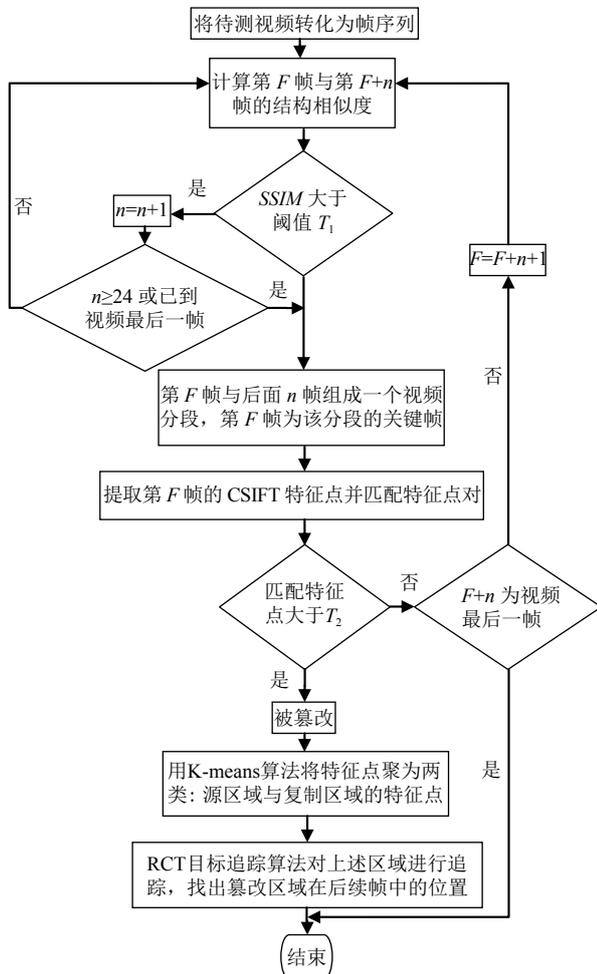


图2 算法流程图

SSIM 的值在 0 到 1 之间分布, 越接近 1 表示越相似, 越接近 0 表示越不相似. 在实际应用中, 将两幅图像分别都分割成若干个 11×11 像素大小的分块, 总共分成 N 块, i 是块号 (i=1~N), 分别计算两幅图像对应位置的两个块的 SSIM 系数, 最后取平均 SSIM 系数:

$$MSSIM = \frac{1}{N} \sum_{i=1}^N SSIM(x_i, y_i) \quad (13)$$

通过上述介绍的 SSIM 指标对原始视频序列进行分段: 从视频第一帧开始, 将当前帧 F 与其后续每一帧计算 SSIM 值, 当第 F 帧与当前帧的 SSIM 值大于阈值 T₁ 时, 当前帧被判定为与 F 帧同一个视频分段; 当小

于 T₁ 时, 当前帧的前一帧被判定为该视频分段的结尾, 当前帧则设定为下一个视频分段的开始. 这里设定每个分段长度不超过 24 帧, 如图 2 算法流程如所示, 当分段长度 n≥24 时, 不论 SSIM 值的大小, 该分段结束.

2.2 利用 CSIFT 检测复制粘贴篡改行为

对关键帧进行检测. 首先提取关键帧 F 的 CSIFT 特征点以及特征向量, 然后进行特征点匹配. 在同一关键帧内, 匹配上的点对少于阈值 T₂ 对, 则认为该帧未经过帧内复制粘贴篡改, 该关键帧代表的分段都未被篡改, 然后对下一个分段的关键帧进行检测, 直到整个视频结束; 如果匹配点对大于等于 T₂ 对, 则认定该帧被篡改过, 进行下一步篡改区域定位.

2.3 K-means 聚类定位复制粘贴区域

对被判定为篡改帧中的特征点进行 K-means 聚类, 分类结果为两个簇, 这两个簇分别代表复制粘贴篡改的源区域和目标区域. 聚类时随机选择一对已经匹配的特征点对作为初始中心, 进行聚类. 虽然 K-means 聚类方法对初始值较为敏感, 但是本文算法在聚类前对特征点进行了匹配, 且本文所用 CSIFT 方法较其它方法更为合理准确, 其误匹配的点对比率很小. 故本文选择已匹配的特征点对作为初始中心.

K-means 是一种经典、快速的聚类算法, 使用该方法能进一步提高检测的整体时间效率.

2.4 追踪篡改区域

本文使用 RCT (Real-time Compressive Tracking, 实时压缩追踪) 方法分别对源区域和篡改区域进行学习, 并追踪这两个区域在视频后续帧中的位置. 因 RCT 是一种单目标跟踪算法, 故我们需要对发现篡改区域的首帧中聚类产生的两个簇——代表源区域和复制粘贴区域的两个簇分别进行追踪. 首先做一个包含簇中所有元素的外接矩形框, 该矩形框即为追踪窗口或目标区域. RCT 算法会对该目标 (正样本) 以及背景 (负样本) 分别进行学习并建立针对该帧的分类器. 然后通过该分类器对后一帧进行分类, 找出目标区域. 接着对这后一帧也进行类似的学习和建模, 利用新的分类器找出其后续帧的目标区域位置, 循环这一操作直至该目标消失或视频结束.

外接矩形框的大小根据簇的大小和簇中特征点所在高斯金字塔的层共同决定, 方法如下:

步骤一. 先对该簇作一个包含簇中所有特征点的最小矩形框, 如图3中的矩形框 ABCD. 中矩形框内的若干点代表特征点, 矩形 EFGH 代表作为初始目标的外接矩形框, IJ 代表矩形 ABCD 的四条边到矩形 EFGH 的四条边的距离.

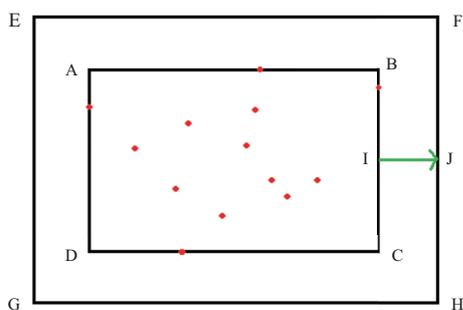


图3 外接矩形框 EFGH

步骤二. 对最小外接矩形框 ABCD 进行向外扩张, 扩张距离 IJ 的长度由该簇中特征点处于高斯金字塔的第几个 octave 决定, 如图4所示, 如果 P 和 Q 是匹配的特征点, 且他们处于高斯金字塔的不同 octave, 如果 Q 处于自上往下第 n 个 octave, 则 Q 所在簇的 IJ 的长度设置为:

$$Lenth_{IJ} = M/n \quad (14)$$

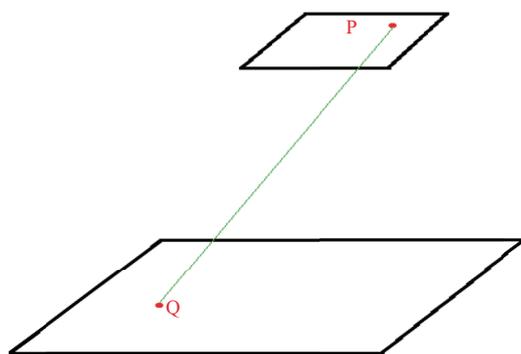


图4 特征点所处不同的 octave

本文实验中, 根据高斯金字塔的 octave 数, 我们设置 M 为 64, $Lenth_{IJ}$ 取值在 4 至 64 这一范围. 外接矩形框 EFGH 即为 RCT 的初始输入, 用于后续跟踪.

RCT 是一种压缩追踪算法, 实时性好, 适合应用于视频检测, 提高了算法整体的运行速率; 另一方面, RCT 也是鲁棒性较高的一种算法, 能抵抗缩放、模糊、抖动、遮挡等多种干扰因素.

RCT 是一种压缩追踪算法, 实时性好, 适合应用于视频检测, 提高了算法整体的运行速率; 另一方面, RCT 也是鲁棒性较高的一种算法, 能抵抗缩放、模糊、抖动、遮挡等多种干扰因素.

3 实验结果

3.1 实验环境及本文方法实验结果

本次实验的视频来源于 SULFA (Surrey University Library for Forensic Analysis) 数据库和其他网络下载, 以及部分用 Canon EOS 6D 相机拍摄的视频, 视频篡改的软件为 Adobe Photoshop CS3 和 Imagineer Systems Mokey V4.1.4. 进行篡改检测的软件环境: Matlab 2013; 硬件环境: 奔腾 E6600 的 CPU, 3 G 内存, Windows 7 操作系统. 实验检测视频数量为 12 个, 总共分为 4 种类型的篡改: 1) 复制区域为静止区域; 2) 复制区域为运动区域; 3) 复制区域进行缩放操作; 4) 复制区域进行了变形扭曲. 针对以上 4 种类型的篡改, 本文算法的部分实验结果如以图5至8所示, 其中红色方形框与红色方形区域代表复制粘贴的源区域与篡改区域.

(1) 复制区域为静止区域



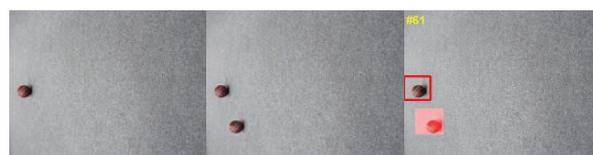
(a) Video1 原始帧 (b) Video1 篡改帧 (c) Video1 检测结果



(d) Video2 原始帧 (e) Video2 篡改帧 (f) Video2 检测结果

图5 复制区域为静止区域的视频的检测结果

(2) 复制区域为运动区域



(a) Video4 原始帧 (b) Video4 篡改帧 (c) Video4 检测结果

图6 复制区域为运动区域的视频的检测结果

(3) 复制区域进行缩放



(a) Video5 原始帧 (b) Video5 篡改帧 (c) Video5 检测结果



(d) Video6 原始帧 (e) Video6 篡改帧 (f) Video6 检测结果

图7 复制区域进行了缩放操作的视频的检测结果



(a) Video7 原始帧 (b) Video7 篡改帧 (c) Video7 检测结果

图8 复制区域进行变形扭曲的视频的检测结果

通过上述实验可以看出本文算法具有一定的鲁棒性,对复制粘贴区域为静止和运动的情况都有很好的检测效果,并且也能较好的检测出复制区域进行了缩放以及变形扭曲的篡改视频.根据本文实验,设定阈值 $T_1=0.97, T_2=5$.

3.2 CSIFT 与 SIFT 特征提取比较

图9和图10分别为CSIFT和SIFT方法在Video1的第267帧上提取和匹配特征点时的效果图,可以看出SIFT方法存在一定的误匹配点对,而CSIFT方法更准确.

3.3 对比实验

实验结果如表1和表2所示,表1为检测的准确率和召回率对比,表2为算法检测总时间的对比.因文献[3,19]不能检测复制粘贴区域旋转缩放等情况,故本文与文献[3,19]的比较只针对于复制粘贴区域完全相同的视频.表2中“/”代表该方法不能检测对应的篡改视频,故没有检测时间;表1中的准确率和召回率也只计算了检测复制粘贴区域完全相同的篡改视频所得到的实验结果.Video1–Video8为进行了复制粘贴篡改的视频,Video9–Video12为未被篡改的视频.

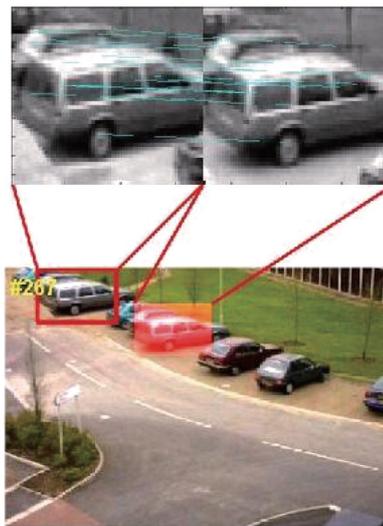


图9 特征点匹配 (CSIFT)

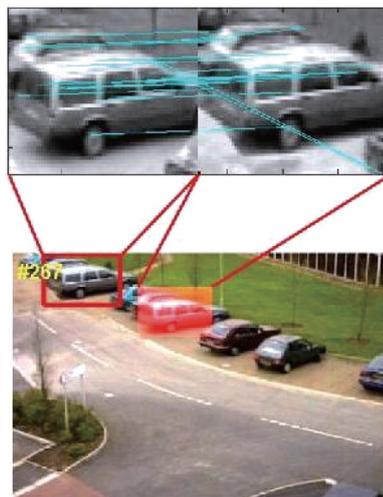


图10 特征点匹配 (SIFT)

为了评估算法性能,本文使用准确率和召回率对实验结果进行分析,准确率 *Precision*, 召回率 *Recall* 定义如下:

$$Precision = \frac{N_c}{N_c + N_f} \tag{15}$$

$$Recall = \frac{N_c}{N_c + N_m} \tag{16}$$

其中, N_c 为正确检测出复制粘贴篡改行为的帧数, N_f 为错误检测的帧数,即原始帧被检测为篡改帧的数目, N_m 为误认为是原始帧的篡改帧数,即漏检的帧数.

通过分析表2可以得出结论,本文与文献[6]算法因先提取了视频的关键帧,并仅对关键帧进行了检测,故在时间效率上具有一定优势,大大提高了效率.而文

献[19]的方法由于对篡改视频帧视频进行了 16×16 重叠分块,并两两比较,故计算量最大、速度最慢,而当待测视频被判定为原始视频时无需分块比较,故速度较快,甚至在部分实验如检测 Video12 时检测速度超过了本文与文献[8].另外在准确率和召回率方面,本文表现较好,文献[3]虽然在召回率上略高于本文算法,但

该方法仅针对没有旋转和缩放造作的复制粘贴篡改视频,所以总体上看,本文算法检测效果最佳.

表1 本文算法与其它文献的准确率和召回率对比(单位: %)

视频	本文算法	文献[3]	文献[8]	文献[19]
<i>Recall</i>	98.7	99.1	97.2	63.7
<i>Precision</i>	100	99.7	95.4	78.9

表2 本文算法与其它文献算法的单帧平均运行时间对比

视频	视频来源	帧幅像素	本文算法(s)	文献[3](s)	文献[8](s)	文献[19](s)
Video1	网络	336×224	0.134 276	0.314 1693	0.143 215	41.982 116
Video2	网络	720×576	0.184 032	0.596 2738	0.127 642	406.701 56
Video3	网络	320×240	0.422 624	0.963 2053	0.467 960	527.060 141
Video4	SULFA 数据库	320×240	0.026 194	0.145 2074	0.062 664	0.362 610
Video5	SULFA 数据库	320×240	0.142 845	/	0.155 253	/
Video6	SULFA 数据库	320×240	0.123 482	/	0.141 492	/
Video7	SULFA 数据库	320×240	0.096 592	/	0.117 824	/
Video8	SULFA 数据库	320×240	0.117 532	0.282 0705	0.121 635	0.409 988
Video9	SULFA 数据库	320×240	0.435 843	1.102 9794	0.329 536	0.437 411
Video10	Canon EOS 6D 拍摄素材	640×480	0.417 036	1.151 0194	0.295 003	0.449 725
Video11	Canon EOS 6D 拍摄素材	1280×720	0.487 863	1.287 7156	0.513 168	8.692 592
Video12	Canon EOS 6D 拍摄素材	1920×1080	1.015 382	2.339 4401	2.102 917	0.319 877
平均时间			0.300 308	0.909 1201	0.403 1755	109.601 780

4 结束语

本文算法使用了保留图像颜色信息的 CSIFT 特征提取方法,算法首先通过比较结构相似度将视频分段并提取每段第一帧为关键帧;然后提取关键帧的 CSIFT 特征点并匹配点对;匹配点对数目大于阈值的关键帧判定为被篡改帧,最后定位篡改帧中的复制粘贴区域,并利用 RCT 目标跟踪算法跟踪篡改区域在关键帧后续的其它帧上的位置.通过实验结果可以看出,算法具有一定的鲁棒性,能够检测经过了旋转、缩放、变形等后续操作的复制粘贴区域. CSIFT 特征提取方法,与传统算法相比速度更快,准确性更高.再结合目标追踪算法,不需要每一帧都提取 CSIFT 特征并定位篡改区域,进一步提高了检测效率.

参考文献

- 卢贺楠,黄添强,林晶.基于动态阈值的帧复制粘贴篡改检测.计算机系统应用,2016,25(12):108-116.[doi:10.15888/j.cnki.csa.005499]
- Su LC, Li CH, Lai YC, et al. A fast forgery detection algorithm based on exponential-Fourier moments for video region duplication. IEEE Transactions on Multimedia, 2017. [doi:10.1109/TMM.2017.2760098]

- Subramanyam AV, Emmanuel S. Video forgery detection using HOG features and compression properties. IEEE 14th International Workshop on Multimedia Signal Processing. Banff, Canada. 2012. 89-94.
- Wang WH, Farid H. Exposing digital forgeries in video by detecting double quantization. Proceedings of the 11th ACM Workshop on Multimedia and Security. Princeton, NJ, USA. 2009. 39-48.
- Chen W, Shi YQ. Detection of double MPEG compression based on first digit statistics. In: Kim HJ, Katzenbeisser S, Ho ATS, eds. Digital Watermarking. Berlin, Heidelberg. Springer. 2009.
- Zhang J, Su YT, Zhang MY. Exposing digital video forgery by ghost shadow artifact. Proceedings of the First ACM Workshop on Multimedia in Forensics. Beijing, China. 2009. 49-54.
- Pandey RC, Singh SK, Shukla KK. Passive forensics in image and video using noise features: A review. Digital Investigation, 2016, (19): 1-28.
- 杜振龙,焦丽鑫,李晓丽,等.帧内复制粘贴视频伪造的盲检测.中国图象图形学报,2014,19(6):825-834.
- 赵洁,郭继昌,张艳,等. JPEG 图像双重压缩偏移量估计的篡改区域自动检测定位.中国图象图形学报,2015,20(10):1304-1312. [doi:10.11834/jig.20151003]

- 10 Warbhe AD, Dharaskar RV, Thakare VM. A survey on keypoint based copy-paste forgery detection techniques. *Procedia Computer Science*, 2016, (78): 61–67. [doi: [10.1016/j.procs.2016.02.011](https://doi.org/10.1016/j.procs.2016.02.011)]
- 11 林晶, 黄添强, 林玲鹏, 等. 采用局部强度顺序模式的图像复制—粘贴篡改检测算法. *通信学报*, 2016, 37(Z1): 132–139. [doi: [10.11959/j.issn.1000-436x.2016259](https://doi.org/10.11959/j.issn.1000-436x.2016259)]
- 12 Abdel-Hakim AE, Farag AA. CSIFT: A sift descriptor with color invariant characteristics. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA. 2006. 1978–1983.
- 13 朱叶, 申铨京, 陈海鹏. 基于彩色 LBP 的隐蔽性复制-粘贴篡改盲鉴别算法. *自动化学报*, 2017, 43(3): 390–397.
- 14 Geusebroek JM, Van den Boomgaard R, Smeulders AWM, *et al.* Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(12): 1338–1350. [doi: [10.1109/34.977559](https://doi.org/10.1109/34.977559)]
- 15 Kubelka P. New contributions to the optics of intensely light-scattering materials. Part II: Nonhomogeneous layers. *Journal of the Optical Society of America*, 1954, 44(4): 330–335. [doi: [10.1364/JOSA.44.000330](https://doi.org/10.1364/JOSA.44.000330)]
- 16 Wyszecki G, Stiles VS, Kelly KL. Color science: Concepts and methods, quantitative data and formulas. *Physics Today*, 1968, 21(6): 83. [doi: [10.1063/1.3035025](https://doi.org/10.1063/1.3035025)]
- 17 Lowe DG. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004: 91–110.
- 18 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612. [doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861)]
- 19 Wang WH, Farid H. Exposing digital forgeries in video by detecting duplication. *Proceedings of the 9th ACM Workshop on Multimedia & Security*. Dallas, TX, USA. 2007. 35–42.