

基于内容的视频取证研究

惠雯^{1,2} 赵海英³ 林闯² 杨扬⁴

(北京科技大学计算机科学与技术系 北京 100083)¹ (清华大学计算机科学与技术系 北京 100084)²
(新疆师范大学计算机科学技术学院 乌鲁木齐 830054)³ (北京科技大学通信工程系 北京 100083)⁴

摘要 基于内容的视频取证技术是检验视频内容合法性的一种有效方法,目前已成为多媒体安全领域的研究热点。该技术利用视频指纹,直接提取媒体对象内容的各种特征作为证据,具有较强的实用性。在典型的视频取证模型中,提取视频指纹需要经过镜头分割、关键帧选取和指纹提取等过程。分析和总结了这3个过程所涉及的各种算法及其近年来的研究进展情况。

关键词 视频取证,视频指纹,镜头,关键帧

中图分类号 TN919.8 **文献标识码** A

Content-based Video Forensics

HUI Wen^{1,2} ZHAO Hai-ying³ LIN Chuang² YANG Yang⁴

(Department of Computer Science and Technology, University of Science and Technology Beijing, Beijing 100083, China)¹

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)²

(College of Computer Science and Technology, Xinjiang Normal University, Urumqi 830054, China)³

(Department of Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China)⁴

Abstract Content-based video forensics is an effective technique to detect the legality of video content. Nowadays, it has been a hot in the research field of multimedia security. By taking advantage of video fingerprint, this technique can extract varieties of visual features from illegal videos as evidence. In a typical video forensics model, the process of fingerprint extraction consists of shot segmentation, key frame extraction and fingerprint extraction. Thus, this paper analyzed content-based video forensics technique from these three aspects, and introduced the research development of the related algorithms.

Keywords Video forensics, Video fingerprint, Shot, Key frame

1 引言

随着数字视频的广泛传播,其内容合法性逐渐成为一个迫切需要解决的问题。借助专业的视频取证技术(digital video forensics)^[1-3]能够可靠而高效地识别视频中包含的盗版或违禁内容,提供有说服力的证据。目前该技术已成为多媒体安全领域的研究热点。

常用的视频取证技术主要有3种:数字签名(digital signature)^[4]、视频水印(video watermark)^[5]和视频指纹(video fingerprint)技术^[6]。其中前两种方法要求内容提供方必须对视频进行预处理,如提取摘要或嵌入验证信息,还要由内容提供方、取证方及可能存在的第三方可信机构,在共享算法和密钥的前提下进行取证,而实际情况往往需要在不依赖任何先验信息和多方参与的前提下检验视频内容。例如,互联网中的视频来源广泛,大都不含数字水印或数字摘要,这就给取

证工作带来了一定的困难。因此,文献[10]研究了视频指纹技术,这是一种基于内容的视频取证技术,借助该技术,取证方可独立完成对视频内容合法性的检验,且无需事先对视频做任何预处理。该技术实用性较强,应用空间较大,是本文综述的主题。

2 基于内容的视频取证理论框架

所谓基于内容的视频取证,是直接提取媒体对象内容的各种特征作为证据。它采用视频指纹技术^[7],自动识别、提取和压缩视频数据。类似于人的指纹对人的作用,视频指纹表示视频内容的内在本属性,可唯一地标识一段视频序列,因此也可称其为“视频DNA”。一个好的视频指纹通常具有如下特性^[6]:

1)鲁棒性(robustness)。对同一媒体的不同变化(失真)具有唯一性,即对同一个媒体内容的多种变形(如不同压缩编

到稿日期:2011-02-21 返修日期:2011-04-19 本文受国家自然科学基金(60873254, 60863010), 973 国家重点基础研究发展计划前期项目(2010CB334709)资助。

惠雯(1982—),女,博士生,主要研究方向为多媒体通信, E-mail: hwen97@126.com; 赵海英(1969—),女,副教授,主要研究方向为模式识别与图像处理; 林闯(1948—),男,教授,博士生导师,主要研究方向为计算机网络与系统性能评价、多媒体通信; 杨扬(1955—),男,教授,博士生导师,主要研究方向为多媒体通信、网络计算、无线通信、图像处理与模式识别。

码格式、模数转换、尺寸变化等),其指纹保持不变。

2)独立性(pairwise independence)。对不同视频具有强区分性,即按照同一算法从不同媒体内容中提取的指纹应该有较大差别。

3)搜索高效性(database search efficiency)。指纹数据量小,且易于做高速匹配搜索。

典型的基于内容的视频取证模型如图1所示。首先收集已知的非法视频,提取其指纹信息,建立非法视频指纹库。这项工作可由权威部门完成,或由取证工作者逐步积累完成。然后,计算目标视频的指纹,并与非法视频指纹库进行匹配,以此判断目标视频的原创性及内容合法性。

在视频取证过程中,一个关键问题就是如何提取视频指纹。典型的视频指纹提取模型如图2所示,其涉及的关键技术包括镜头分割、关键帧选取和视频指纹的提取,下面将分别进行论述。

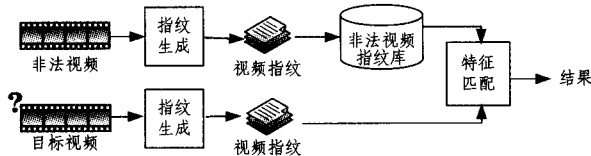


图1 基于内容的视频取证模型

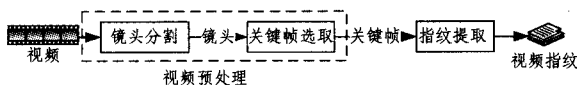


图2 指纹提取模型

3 镜头分割

视频中包含丰富的信息,因此无法像文本一样直接描述或比较其内容。要获取视频内容,首先需要对其结构进行分析。视频从结构上自顶向下可分为视频(video)、场景(scene)、镜头(shot)和帧(frame)4个层次,如图3所示。其中,帧是视频数据的最小单元,是一个静止的画面;镜头是一段中间没有摄像机信号中断、连续拍摄的画面,它代表一个场景中在时间和空间上的连续动作,是视频数据的基本单位,也是构成视觉语言的基本单位;场景由内容相近的镜头组成,从不同角度描述同一个事件;而视频则由多个场景构成,表示一个完整事件。

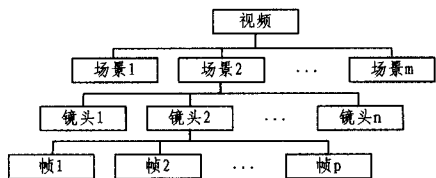


图3 视频数据的结构

将视频自动分割为镜头,再在镜头的基础上提取视频指纹,可简化指纹提取的复杂度,提高效率。镜头检测即找到镜头与镜头之间的切换点。有的视频切换少,每个镜头持续时间较长,如体育节目;而有的视频则切换频繁,镜头持续时间较短,如新闻节目。一般认为镜头的切换方式主要分为两类^[8]:渐变(gradual change)和突变(abrupt change)。突变是由前一个镜头直接转换为下一个镜头,中间没有时间上的延迟,常在两帧间完成;渐变则是加入了一些空间或时间上的编辑效果,由前一个镜头逐渐转换为下一个镜头,常延续十几到

几十帧。渐变的方式有很多,常见的有淡化(fade)、溶解(dissolve)和扫换(wipe)等。淡化是画面逐渐加强(淡入, fade in)或画面慢慢减弱至消失(淡出, fade out);溶解是前一个镜头画面逐渐减弱,同时下一个镜头画面逐渐加强;扫换则是从画面的某一部分开始,前一个镜头逐渐被下一个镜头代替。无论哪种切换,在切换的位置,视频数据都会发生一系列变化,通常表现为颜色差异突然增加、新旧边缘的远离、对象形状的改变和运动的不连续等。根据这些变化规律,就可以定位镜头边界。

检测镜头边界常用的方法包括模板匹配法、直方图法、边缘法和模型法。

3.1 模板匹配法

模板匹配法^[8]又称像素比较法。该方法将两帧对应像素的灰度或颜色差的绝对值之和作为帧间差。若大于某阈值,则认为有镜头切换。如RGB图像的帧间差为

$$d_{RGB}(I_i, I_j) = \sum_k (|H_i^r(k) - H_j^r(k)| + |H_i^g(k) - H_j^g(k)| + |H_i^b(k) - H_j^b(k)|) \quad (1)$$

式中, H 为直方图的值。

这种方法原理简单、便于实现,对检测镜头突变很有效。但对于噪声和运动物体,方法本身与像素位置密切相关,因此可能计算出大于阈值的帧间差,造成误判。对此,Zhang等^[11]提出了一种改进方法,即将各帧划分为 8×8 的像素块,对每块取灰度或颜色的均值,再根据这个均值对比相邻两帧的对应块。这样可减小物体运动和噪声的影响。

3.2 直方图法

直方图法^[12]是最常用的计算帧间差的方法。它不考虑像素的位置信息,而是利用灰度或颜色的统计值,将颜色空间划分为若干个离散的小区域,计算落入每个小区域中的像素数。由直方图的相似度衡量图像之间的相似情况,以此判断是否有镜头切换。设颜色空间分为 n 个区间, H_{ik} 是第 i 帧中落入第 k 个颜色区间的像素数,则帧间差为

$$d(I_i, I_j) = \sum_{k=1}^n |H_{ik} - H_{jk}| \quad (2)$$

这种方法不考虑像素的位置信息,因而抗噪声能力较强,且适合检测镜头突变。不足之处是,有时会漏掉镜头变换,因为两幅结构完全不同的图像的直方图也可能相近;还有对视频序列中亮度的突然变化,容易产生误判。

Nagasaka和Tanaka^[13]提出了一种将视频帧划分为 4×4 相同大小的子块并比较相应子块的方法。还有一种类似于直方图法的 χ^2 直方图法^[8],相关研究文献^[14,15]证明其检测效果要优于模板匹配法和直方图法。

3.3 边缘法

边缘法的基本思想是“在发生镜头转换时,新出现的边缘应远离旧边缘的位置,同样旧边缘消失的位置应远离新边缘的位置”^[20]。提取相邻两帧视频图像的边缘,计算出现的新边缘与消失的旧边缘的像素比例。当二者中的较大值大于某阈值时,认为有镜头切换。该方法适合检测镜头突变;但计算量较大,且当边缘不明显时,识别效果不够理想。

以上3种方法都是通过帧间差来定位镜头边界,适用于突变切换的检测,因为在镜头发生突变时,其切换处的视频帧会有明显变化。而对于渐变,其帧间差虽有所改变,但变化比较平缓,很容易导致误判。为区分渐变切换与镜头的运动,

Zhang等^[11]采用了一种光流检测法。光流计算常用于运动目标跟踪,它通过相邻两帧中存在的空间局部性原理进行速度场分析,可较好地反映出渐变切换的位置。但计算过程非常复杂,且在镜头间颜色直方图很接近或光照变化很大的情况下会失效。

3.4 模型法

对于渐变镜头,整个变换可看作一个连续过程,可近似认为该过程符合某个线性直线方程,所以有学者提出了模型法^[21]。这是一种基于镜头编辑时采用渐变效果的数学模型的方法,它需要针对不同的渐变效果,建立不同的数学模型,因此具有很大的局限性,建模过程也比较复杂;还有,该方法对运动和噪声十分敏感,只有在含有少量运动的视频序列中才有比较好的效果。

此外,为提高检测效率,还有人提出了压缩域法^[23]。这是一个较新的研究方向,它针对以压缩形式保存的视频数据,直接利用DCT系数、运动向量及运动补偿宏块等对压缩过程中已提取的信息进行镜头边界检测,大大降低了计算复杂度,节省了解码时间。Yeo和Liu^[24]的研究证明了直接在压缩域上计算的帧间差的分布与解压后帧间差的分布的对应关系。

表1对上述4种主要的镜头分割方法进行了分类比较。总体来讲,检测突变的方法都比较成熟,成功率也较高;最大的难点在于检测渐变,现有方法只局限于简单渐变的检测,还没有一种对不同渐变类型都有较好效果的方法。这一问题仍有待进一步研究。

表1 主要镜头分割方法比较

算法	计算复杂度	精度	效率
模板匹配法 (像素比较法)	低	对镜头突变精度较高,但对运动和噪声很敏感	高
直方图法	较低	对镜头突变精度较高,抗噪能力较强,对镜头内全局运动敏感、局部运动不敏感	高
边缘法	较高	对镜头突变和画面简单的视频精度高,光线较暗和轮廓不明显时鲁棒性较差,对快速运动的物体很敏感	较低
模型法	较高	对镜头渐变精度高,但有很大局限性,对运动和噪声很敏感	较低

4 关键帧选取

镜头中的各个视频帧不是同等重要的,一些帧可能反映更多有关对象、运动和事件的信息,这些帧即为关键帧。去掉冗余数据,保留关键帧,在此基础上计算视频指纹,既可简化计算过程,又不会损失精度。选取关键帧遵循的原则是,应尽可能准确地描述镜头中的主要事件,且数据量应尽量小,计算过程应尽量简单。

关键帧选取分为静态和动态两种。静态关键帧选取一般只适合于静止的镜头;动态关键帧选取则是根据镜头内视频帧之间的变化程度,动态地将镜头分割成分镜头,进而对每个分镜头选取关键帧。具体选取方法有很多,主要可分为以下4类。

4.1 基于镜头法

早期选取关键帧的尝试主要是利用颜色特征。一个最简单的方法是将镜头第一帧、中间帧或最后一帧选作关键帧^[25],称为首(尾)帧法。当然,这样选出的关键帧很不准确,

只适合于内容变化小或基本不变的镜头。对于摄像机不断运动的镜头,该方法不稳定,无法有效描述镜头的主要内容。

除了在镜头内固定位置选取关键帧外,基于镜头法中也包括通过计算平均值选取关键帧的方法。如帧平均法,计算镜头内所有帧在某个位置上的像素平均值,若某帧在该位置上的像素值接近这个平均值,即为关键帧;直方图平均法^[11],对镜头内所有帧的统计直方图取平均,再选择与该平均直方图最接近的帧作为关键帧。这类方法最大的优点是计算简单,所选帧具有平均代表意义;但由于只从一个镜头中选取一个关键帧,因此无法描述有多个物体运动的镜头。

总体上讲,从镜头中选择固定数量的关键帧不是一种好方法,它有可能造成一些简单镜头中产生过多相似的关键帧,而一些复杂镜头中的关键帧又太少,遗漏了比较重要的片段。

4.2 基于图像特征法

按照信息论的原理,不同图像要比相似图像携带更多的信息。也就是说,将最不相关的帧选为关键帧,才能包含最多的信息。这样,利用颜色、纹理的改变^[26],或利用互信息量^[27]作为相似度联合聚类算法计算视觉内容的变化,再根据这些信息发生变化的显著程度即可确定关键帧。如先把镜头第一帧定为关键帧,然后计算第一帧与剩余帧之差。当大于某一阈值时,就选当前帧为关键帧,并继续计算当前帧与剩余帧之差。重复这一过程,直至检测到最后一帧。这种算法还产生了一系列变型^[28],其主要区别是在距离度量上。

这类方法是根据镜头内容的变化程度选取相应数目的关键帧,难点是阈值的选择;问题主要在于特征比较单一,所选帧未必真正具备代表性,而且当有物体快速运动时,关键帧的数目容易过多。

4.3 基于运动分析法

前两类方法都没有考虑运动特征,基于运动分析法可通过分析相机运动^[29]、光流、MPEG-7运动描述符^[30]等选取关键帧。常用的是光流分析法^[31],这是一种根据镜头结构选择相应数目关键帧的方法,其基本思想是视频数据中的静止或停留意味着强调某一动作,所以通过光流分析计算出镜头中的运动量,在运动量局部最小值处即可取为关键帧。如果先从背景中分离出运动对象,再计算对象所在位置的光流,则效果更好。

不足之处是,这类方法的计算量都非常大,而且局部最小值也不一定准确;依赖于局部信息,鲁棒性不强;没有足够重视由累加动态带来的内容变化。一种改进方法^[32]是利用运动感知能量(Perceived Motion Energy, PME)模型对视频中的运动模式进行建模,处于加速或减速运动的帧被选为关键帧。

4.4 基于聚类分析法

Song和Fan^[33]利用一种被广泛应用在模式识别、语音分析和信息检索中的聚类算法选取关键帧。首先初始化一个聚类中心,然后根据当前帧与中心的距离判断是归入该类还是作为新的聚类中心,最后将各类中离聚类中心最近的帧当作关键帧。这种方法同样有选择阈值的困难,虽然二次聚类^[34]可自适应地得到关键帧聚类的阈值,但时间代价非常大。一种改进方法是基于直方图的模糊c-均值聚类算法(HBF-CM)^[35],这种方法的优点是无需预先设定阈值。Zeng等^[36]提出了一种优势集聚类算法,其根据视频镜头的复杂度为镜

头选择最优聚类数,可减小计算复杂度,并且为保证关键帧不会在时间上距离太近,将序列中间帧选作关键帧。Chasanis等^[37]采用一种改进的谱聚类算法,即通过相似矩阵的特征值来计算关键帧的数量。Panagiotakis等^[9]提出一种基于计算几何的算法,即将视频序列抽象为一条“内容曲线”,通过平均分割内容曲线获得关键帧。这样可保证关键帧准确描述视频各部分的内容,且彼此间等价。

此外,还有全局最优估计法,它以颜色作为特征相似度,对视频序列进行全局处理,如通过计算像素最大出现概率和图像帧序列的最小生成树来提取关键帧^[38]。Song和Fan^[39]将选择关键帧与提取视频对象相结合,通过挖掘二者之间的内在联系,利用高斯混合模型(Gaussian Mixture Model, GMM)较好地解决了关键帧选择的问题。

表2对上述4种主要的帧选取方法进行了比较。

表2 主要关键帧选取方法比较

算法	计算复杂度	精度	效率
基于镜头法	低	适合于内容变化小的镜头,对复杂镜头精度较低	高
基于图像特征法	低	精度受阈值影响较大	较高
基于运动分析法	高	对运动特征明显的镜头精度较高	较低
基于聚类分析法	低	精度受阈值影响较大	较低

5 媒体指纹提取

视频指纹提取可基于任何视觉特征,包括(但不限于)视频流中的关键帧序列分析、色彩和运动变化等特征。它根据视频本身的感知性来识别视频内容,不受视频编码及一些数字处理(如改变亮度、分辨率、帧格式等)的影响。相比源文件,指纹的数据量要少很多,能够比较容易地存储在数据库中用于查询匹配,但不能用来重建原始视频内容。

视频指纹中所采用的算法早在1998年就有人在文献^[40,41]中提出,用于立体图像匹配。由于视频是由许多幅按时间顺序排列的连续图像(即视频帧)组成,因此后来产生的很多算法都是源于图像处理中的方法。当然,利用视频独有的运动特征,也可以提取指纹。根据所选特征的不同,可将指纹提取方法分为两大类:静态特征提取方法和运动特征提取方法。

5.1 静态特征提取方法

静态特征的提取主要是针对关键帧,可采用一些图像处理中的算法,如基于关键帧的颜色、纹理、形状、边缘等特征提取指纹,其中颜色和纹理特征用得最多。

序数度量法是一种比较有代表性的方法。它源于Bhat和Nayar^[40]在立体图像匹配中所提出的方法,Mohan^[41]将该方法应用于视频序列匹配,此后在不同的算法中都用到了该方法^[42,43]。其主要思想是将视频帧分为 N 个大小相等的块,求出每一块的亮度均值 I_i ($1 \leq i \leq N$),然后按均值大小给每一块分配序号 r_i ,最后按顺序把序号排列成一个 N 维矢量:

$$F(t) = (r_1, r_2, \dots, r_N) \quad (3)$$

则两部视频 R 和 C 在 t 时刻的距离为

$$D(t) = \frac{1}{T} \sum_{i=-T/2}^{T/2} |F_R(i) - F_C(i)| \quad (4)$$

式中, T 为一个设定的时间段。

序数度量特征最大的优点就是对于图像质量变化和光照强度的变化不敏感,所以对压缩和误码失真具有非常好的鲁

棒性。这一特性已得到广大学者的充分重视,并在此基础上提出了多种改进算法^[44,45]。此外,该方法的计算过程比较简单,便于实现。其不足之处是,当 N 取值较小时,提取的指纹唯一性较差;还有对图标插入和几何变换等操作鲁棒性也很差。

颜色直方图法是提取颜色特征的经典方法,具有与生俱来的旋转不变性、尺度不变性和平移不变性,近年来在很多算法中都融入了颜色直方图的思想^[12]。但这种方法有一个缺陷,当图像中的颜色不能取遍所有可取值时,直方图中会出现一些零值,对匹配算法的正确性造成影响。累加直方图可减少零值的数量,使两种颜色在特征轴上的距离与它们的相似度成正比,还能克服直方图量化过粗或过细造成的匹配精度下降。Yang等^[46]针对颜色直方图法提出了一种改进算法,把一个镜头分为 k 个大小相等的段,计算每段在时间轴上分布的各帧对应像素点的均值,将其合成一幅图像来表示该段,然后将每幅合成图像均分为块,从每块中提取主颜色和次颜色,这样每个镜头就能表示为一个固定长度的字符串。该算法对由帧速率改变引起的噪声和丢帧现象具有较好的鲁棒性。

纹理特征分析方法大致可分为结构型和统计型两类。其中结构型方法适用于规则纹理,在一般图像中很少见;统计型方法则适合于准规则纹理,即纹理基元没有明确的外形,只是某种灰度或颜色的分布。灰度共生矩阵(GLCM)已被证明是一种很好的统计型纹理特征分析方法,但由于GLCM是像素距离和角度的矩阵函数,因此完整的GLCM计算,其参数选取范围很广,计算量很大。为解决这一问题,薄华等^[47]利用马尔可夫链的性质,从理论上证明了GLCM的计算无需选取所有参数,也可完整描述图像的纹理特征。心理学特征法等也可用于统计型纹理特征的提取。另外,还有基于信号处理的纹理分析法,如小波变换^[14]、Gabor滤波器^[16]以及基于模型的纹理分析法,如马尔可夫随机场模型^[17]、Wold模型^[18]等。

形状特征主要包括基于边界的形状描述和基于区域的形状描述。其中基于边界的形状描述中,最为成功的是傅里叶描述子^[19],它对边界进行傅里叶变换,将其作为形状特征描述;基于区域的形状描述主要有各种矩不变量及各种简单的形状因子^[22],如面积、圆度、偏心度、主轴方向等。为解决遮挡问题,还常使用一些局部特征,如直线段、圆弧、角点、高曲率点等。

5.2 运动特征提取方法

视频数据除了具有静态特征外,还具有运动特征,它反映了视频数据的时域变化。摄像机的运动往往会给视频图像带来全局的影响,如水平移动镜头会使所有像素点也水平移动,焦距拉长会使像素点从中心向四周发散,焦距缩短会使像素点从四周向中心会聚。在只有对象运动时,大部分背景像素不变,而只是运动对象和被遮挡的部分发生变化。由于运动特征无法从一幅静止图像中获得,因此必须分析视频序列。

一种典型算法是基于宏块运动向量的算法^[48]。该算法计算MPEG视频流中 B 帧和 P 帧的运动向量在8个方向上的分布,即

$$F_j = \frac{1}{K} \sum_{i=1}^K \theta(i, j) \quad (5)$$

式中, F_j 表示第 j 个方向的运动向量所占的比例; K 是帧中宏块的总数; 而 $\theta(i, j)$ 取值为 0 或 1, 表示第 i 个运动向量是否在方向 j 上。加上没有运动的块所占的比例, 就得到一个有 9 个分量的特征向量, 即视频指纹。

还有一类运动流算法^[49], 其主要思想是只记录两个连续帧之间相关宏块的运动方向和大小(运动矢量只包含局部数据而没有更多的语义信息)。利用两个连续帧之间运动方向、颜色分布和宏块重叠区域的一致性, 将所有近邻的运动矢量“链接”起来, 这些链接起来的运动向量即为“运动流”特征。

结束语 视频取证是一个迅速成长的研究领域, 在保护多媒体信息安全方面具有十分重要的应用价值。通过以上对视频取证的原理和相关算法的探讨, 可以看出, 经过这几年的发展, 该技术已在理论和实践上取得不少成绩。但随着计算机与网络技术的发展, 视频取证必定面临新的挑战。综合起来看, 我们认为视频取证可从以下几方面进行进一步研究:

第一, 在典型的视频取证模型的基础上, 将视频处理过程简化, 脱离镜头、关键帧等概念, 如利用视频数据的动态性和实时性等特点, 根据视频中的运动对象特征直接提取指纹。

第二, 在复杂场景下, 误检率往往很高。结合对视频图像语义的理解和对视频序列中显著对象的运动分析, 将是另一个发展方向。

第三, 在保证取证精度的同时, 效率也是一个不可忽视的问题, 特别是对大量视频数据的处理, 如对互联网视频进行检测会耗费相当长的时间, 因此有必要研究适合大规模视频取证的算法。

参 考 文 献

[1] Wang Wei-hong, Farid H. Exposing digital forgeries in video by detecting double mpeg compression [C]//Proceedings of the 8th Workshop on Multimedia and Security. Geneva, Switzerland, 2006; 37-47

[2] Wang Wei-hong, Farid H. Exposing digital forgeries in interlaced and deinterlaced video [J]. IEEE Transactions on Information Forensics and Security, 2007, 2(3): 438-449

[3] Wang Wei-hong, Farid H. Exposing digital forgeries in video by detecting duplication [C]//Proceedings of the 9th Workshop on Multimedia and Security. Dallas, Texas, USA, 2007; 35-42

[4] Ahmed F, Siyal M. A robust and secure signature scheme for video authentication [C]//Proceedings of the IEEE International Conference on Multimedia and Expo. 2007; 2126-2129

[5] Lee M J, Kim K S, Suh Y H, et al. Improved watermark detection robust to camcorder capture based on quadrangle estimation [C]//Proceedings of the 16th IEEE International Conference on Image Processing. 2010; 101-104

[6] Lee S, Yoo C. Robust video fingerprinting for content-based video identification [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(7): 983-988

[7] Lee S, Yoo C. Robust video fingerprinting based on affine covariant regions [C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Las Vegas, Nevada, USA, 2008; 1237-1240

[8] Jiang Hai-tao, Helal A, Elmagarmid A K, et al. Scene change detection techniques for video database systems [J]. Multimedia Systems, 1998, 6(3): 186-195

[9] Panagiotakis C, Doulamis A, Tziritas G. Equivalent key frames

selection based on iso-content principles [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(3): 447-451

[10] Yang Xian-feng, Sun Qi-bin, Tian Qi. Content-based video identification; a survey [C]//Proceedings of the International Conference on Information Technology; Research and Education. 2004; 50-54

[11] Zhang Hong-jiang, Wu Jian-hua, Zhong Di, et al. An integrated system for content-based video retrieval and browsing [J]. Pattern Recognition, 1997, 30(4): 643-658

[12] Liu Lu, Lai Wei, Hua Xian-sheng, et al. Video histogram: A novel video signature for efficient web video duplicate detection [J]. Advances in Multimedia Modeling, 2006; 94-103

[13] Nagasaka A, Tanaka Y. Automatic video indexing and full-video search for object appearances [J]. Journal of Information Processing, 1992, 15(2): 316

[14] Vo A, Oraintara S. A study of relative phase in complex wavelet domain; Property, statistics and applications in texture image retrieval and segmentation [J]. Signal Processing; Image Communication, 2010, 25(1): 28-46

[15] Zhang Hong-jiang, Kankanhali A, Smoliar S W. Automatic partitioning of full-motion video [J]. Readings in multimedia computing and networking, 2002; 321

[16] 赵海英, 冯月萍. 应用 Gabor 滤波器和局部边缘概率直方图的全局纹理方向性度量 [J]. 光学精密工程, 2010, 18(7): 1668-1674

[17] You D, Antani S, Demner-Fushman D, et al. Biomedical article retrieval using multimodal features and image annotations in region-based CBIR [C]//Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2010, 7534: 30

[18] 李杰, 朱维乐, 王磊, 等. 基于 Wold 模型和支持向量机的纹理识别 [J]. 计算机研究与发展, 2007, 44(3): 460-466

[19] El-ghazal A, Basir O, Belkasim S. Farthest point distance; A new shape signature for Fourier descriptors [J]. Signal Processing; Image Communication, 2009, 24(7): 572-586

[20] 朱曦, 林行刚. 视频镜头时域分割方法的研究 [J]. 计算机学报, 2004, 27(8): 1027-1035

[21] Filip J, Haindl M. Fast and reliable PCA-based temporal segmentation of video sequences [C]//Proceedings of the 19th International Conference on Pattern Recognition. 2009; 1-4

[22] Li Shan, Lee M C, Pun C M. Complex Zernike moments features for shape-based image retrieval [J]. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 2008, 39(1): 227-237

[23] Patel N V, Sethi I K. Video shot detection and characterization for video databases [J]. Pattern Recognition, 1997, 30(4): 583-592

[24] Yeo B L, Liu B. Rapid scene analysis on compressed video [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 5(6): 533-544

[25] Shahraray B, Gibbon D C. Automatic generation of pictorial transcripts of video programs [C]//Proceedings of SPIE. 1995, 2417; 512-518

[26] Ali A, Mahmood F. Hierarchical Keyframe-based Video Summarization Using QR-Decomposition and Modified k-Means Clustering [J]. EURASIP Journal on Advances in Signal Processing, 2010

Research. 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer Verlag, 2005; 19-35

- [3] 刘建东,余有明,江慧娜. 单向 Hash 函数 SHA-1 的统计分析与算法改进[J]. 计算机科学, 2009, 36(10): 141-145
- [4] National Institute of Standards and Technology (NIST). Cryptographic Hash Algorithm Competition[EB/OL]. <http://csrc.nist.gov/groups/ST/hash/sha-3/index.html>, 2007-11-02
- [5] National Institute of Standards and Technology (NIST). Final Round Candidates[EB/OL]. http://csrc.nist.gov/groups/ST/hash/sha-3/Round3/submissions_rnd3.html, 2009-07-16
- [6] 薛宇, 吴文玲, 王张宜. SHA-3 杂凑密码候选算法简评[J]. 中国科学院研究生院学报, 2009, 26(5): 577-586
- [7] Long M. Implementing Skein Hash Function on Xilinx Virtex-5 FPGA Platform[EB/OL]. http://www.skein-hash.info/sites/default/files/skein_fpga.pdf, 2009
- [8] Namin A H, Hasan M A. Implementation of the Compression Function for Selected SHA-3 Candidates on FPGA[EB/OL]. http://comsec.uwaterloo.ca/seminarfiles/ReviewSeminar2010/Implementation_SHA3_Candidates_on_FPGA.pdf, 2010
- [9] Homsirikamol E, Rogawski M, Gaj K. Comparing . Hardware Performance of Fourteen Round Two SHA-3 Candidates Using FPGAs[EB/OL]. <http://eprint.iacr.org/2010/445.pdf>, 2010
- [10] Baldwin B, Hanley N, Hamilton M, et al. FPGA Implementations of the Round Two SHA-3 Candidates[C]//Field Programmable Logic and Applications (FPL), 2010. Milano, Italy, 2010: 400-407
- [11] Tillich S, Feldhofer M, Kirschbaum M, et al. High-Speed Hardware Implementations of BLAKE, Blue Midnight Wish, Cube-Hash, ECHO, Fugue, Grostl, Hamsi, JH, Keccak, Lua, Shabal, SHAvite-3, SIMD, and Skein[EB/OL]. <http://eprint.iacr.org/2009/510.pdf>, 2009
- [12] Tillich S. Hardware Implementation of the SHA-3 Candidate Skein[DB/OL]. <http://eprint.iacr.org/2009/159.pdf>, 2009
- [13] Hou Fang-yong, He Hong-jun, Xiao Nong. Hash Tree Based Integrity Protection Appropriate for Disk [C]// 2009 WASE International Conference on Information Engineering. Taiyuan, China, 2009; 242-245
- [14] Ferguson N, Lucks S, Schneier B, et al. The Skein Hash Function Family[EB/OL]. <http://www.skein-hash.info/sites/default/files/skein1.3.pdf>, 2010
- [15] Schorr A, Lukowiak M. Skein Tree Hashing on FPGA[C]// International Conference on Reconfigurable Computing. 2010. Cancun, Mexico, 2010; 292-297

(上接第 31 页)

- [27] Cernekova Z, Pitas I, Nikou C. Information theory-based shot cut/fade detection and video summarization [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(1): 82-91
- [28] Cao Yu, Tavanapong W, Kim K. Audio-Assisted Shot Clustering Techniques for Story Browsing. 2008
- [29] 庄越挺, 潘云鹤, 吴飞. 网上多媒体信息分析与检索 [M]. 北京: 清华大学出版社, 2002
- [30] Herranz L, Tiburzi F, Bescos J. Extraction of Motion Activity from Scalable-coded Video Sequences [J]. Semantic Multimedia, 2006; 148-158
- [31] Wolf W. Key frame selection by motion analysis [C]// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. 1996, 2; 1228-1231
- [32] Liu Tian-ming, Zhang Hong-jiang, Qi Fei-hu. A novel video key-frame-extraction algorithm based on perceived motion energy model [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(10): 1006-1013
- [33] Song Xiao-mu, Fan Guo-liang. Joint key-frame extraction and object segmentation for content-based video analysis [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(7): 904-914
- [34] 王方石, 须德, 吴伟鑫. 基于自适应阈值的自动提取关键帧的聚类算法 [J]. 计算机研究与发展, 2005, 42(10): 1752-1757
- [35] Lo C C, Wang S J. Video segmentation using a histogram-based fuzzy c-means clustering algorithm [J]. Computer Standards & Interfaces, 2001, 23(5): 429-438
- [36] Zeng Xianglin, Hu Weiming, Li Wanqing, et al. Key-frame extraction using dominant-set clustering [C]// Proceedings of the IEEE International Conference on Multimedia and Expo. 2008; 1285-1288
- [37] Chasanis V T, Likas A C, Galatsanos N P. Scene detection in videos using shot clustering and sequence alignment [J]. IEEE Transactions on Multimedia, 2009, 11(1): 89-100
- [38] Sze K W, Lam K M, Qiu Guoping. A new key frame representation for video segment retrieval [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(9): 1148-1155
- [39] Song Xiao-mu, Fan Guo-liang. Key-frame extraction for object-based video segmentation [C]// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. 2005; 689-692
- [40] Bhat D N, Nayar S K. Ordinal measures for image correspondence [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(4): 415-423
- [41] Mohan R. Video sequence matching [C]// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. 1998, 6; 3697-3700
- [42] Hua Xian-sheng, Chen Xian, Zhang Hong-jiang. Robust video signature based on ordinal measure [C]// Proceedings of the International Conference on Image Processing. 2005, 1; 685-688
- [43] Kim C, Vasudev B. Spatiotemporal sequence matching for efficient video copy detection [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(1): 127-132
- [44] Paisitkiangkrai S, Mei Tao, Zhang Jian, et al. Scalable clip-based near-duplicate video detection with ordinal measure [C]// Proceedings of the ACM International Conference on Image and Video Retrieval. 2010; 121-128
- [45] Shang Li-feng, Yang Lin-jun, Wang Fei, et al. Real-time large scale near-duplicate web video retrieval [C]// Proceedings of the International Conference on Multimedia. 2010; 531-540
- [46] Yang Xian-feng, Tian Qi, Chang E C. A color fingerprint of video shot for content identification [C]// Proceedings of the ACM International Conference on Multimedia. 2004; 276-279
- [47] 薄华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析 [J]. 电子学报, 2006, 34(1): 155-158
- [48] Manerba F, Benois-Pineau J, Leonardi R, et al. Multiple moving object detection for fast video content description in compressed domain [J]. EURASIP Journal on Advances in Signal Processing, 2008; 5
- [49] Su C W, Liao H Y M, Tyan H R, et al. Motion flow-based video retrieval [J]. IEEE Transactions on Multimedia, 2007, 9(6): 1193-1201