论 文

抗几何失真的视频 Hashing 算法研究

项世军^{①④}, 杨建权^②, 黄继武^{③*}

① 暨南大学信息科学技术学院, 广州 510632

② 中国科学院深圳先进技术研究院, 深圳 518055

③ 中山大学信息科学与技术学院, 广州 510275

④ 信息安全国家重点实验室 (中国科学院信息工程研究所), 北京 100029

* 通信作者. E-mail: isshjw@mail.sysu.edu.cn

收稿日期: 2010-10-07; 接受日期: 2011-04-04

国家自然科学基金 (批准号: 60903177, U1135001)、国家重点基础研究发展计划 (批准号: 2011CB302204)、中央高校基本科研业务费专项资金 (批准号: 21611408) 和中国科学院战略性先导科技专项课题 (批准号: XDA06030600) 资助项目

摘要 文中提出了一种基于视频灰度直方图形状的 Hashing 算法, 能有效抵抗各种常见的几何失真和视频处理操作. 算法的鲁棒性原理如下: 1) 由于直方图的形状与像素位置无关, 故基于直方图的视频 Hashing 算法能有效抵抗各种常见的几何攻击; 2) 由于在计算 Hashing 前对视频帧进行了平滑预处理, 故算法对加噪攻击、模糊滤波、有损压缩等处理操作有很好的鲁棒性; 3) 由于计算 Hashing 前在时间轴上进行了低通滤波预处理, 故算法能抵抗帧率变化、帧丢失等时域同步攻击. 实验结果表明, 所提出的 Hashing 算法有良好的唯一性和鲁棒性能.

关键词 视频 Hashing 几何失真 直方图 统计特征 高斯滤波

1 引言

随着多媒体技术和网络通信技术的发展, 数字视频作品的创作周期越来越短, 作品的分发、传输更多在网络上进行. 如何实现视频作品的版权保护, 视频内容的真实性认证, 视频数据的高效检索成为亟待解决的课题. 视频水印和视频 Hashing 是两种可用于保护视频数据的潜在技术手段. 视频数字水印技术通过在视频数据嵌入版权信息来实现对作品的流通跟踪及拷贝控制. 视频 Hashing 技术无需向作品嵌入认证信息, 没有水印中的嵌入失真问题. 认证时只需将待认证作品的 Hashing 签名与受保护作品数据库中的 Hashing 签名一一进行比对, 一旦发现相同或相似的 Hashing 签名存在, 则认为该视频作品是受保护的, 从而明确版权归属. 视频 Hashing 技术可用于: 1) 认证视频在传输过程中的内容完整性和真实性; 2) 构建基于内容的视频检索系统; 3) 作为水印信号用于设计基于内容的数字水印, 以抵抗稳健水印中的二义性攻击和脆弱水印中的拷贝攻击. 视频 Hashing 技术由于其广阔的应用前景, 受到研究者的重视.

视频 Hashing 是视频数据的一种紧凑表达. 视频 Hashing 函数根据视频数据作用于人眼的感知特性, 生成代表视频内容的固定长度的简短二进制位串. 理想的感知 Hashing 函数当用于基于内容的检索时, 要求具有良好的唯一性和鲁棒性; 当用于内容的真实性认证时, 则侧重于安全性和对篡改的敏

引用格式: 项世军, 杨建权, 黄继武. 抗几何失真的视频 Hashing 算法研究. 中国科学: 信息科学, 2012, 42: 578-587

感性.

目前, 各国学者已提出一些鲁棒的视频感知 Hashing 算法. 文献 [1] 提出一种基于累计直方图量化的针对缓变视频的 Hashing 算法, 对噪声干扰、直方图均衡、伽马校正等内容保持操作具有鲁棒性. 通过 Harris 检测器 [2], 文献 [3] 提取视频帧的特征角点, 并以这些特征点形成的轨迹为特征来构造 Hashing 函数, 对视频的对比度变化、噪声干扰等处理鲁棒. 文献 [4] 采用相邻帧对应块间的相似性来设计 Hashing 函数, 对直方图均衡等一些帧内灰度处理操作鲁棒, 同时能抵抗对视频的 MPEG2 压缩. 文献 [5] 则以视频帧灰度梯度方向的质心为特征来构造 Hashing 函数. 由于灰度梯度的方向取决于灰度的变化而非具体的灰度值, 因此该算法能有效抵抗对视频的全局灰度改变操作. 同时, 文献 [5] 的实验数据表明该算法对帧丢失、帧内裁剪、视频压缩等处理有鲁棒性. 除上述基于视频时空域的 Hashing 方法外, 还有基于变换域的视频 Hashing 算法. 文献 [6] 对预处理后的视频灰度帧集实施 3D-DCT 变换, 并将低频系数量化成 Hashing 签名. 由于 3D-DCT 把视频的时空能量集中到几个低频系数上, 因此, 该算法对常见的视频帧内处理操作和时间轴操作都有较好的鲁棒性. 文献 [7] 提出一种基于 3D-DWT 的视频 Hashing 算法, 并用于广播监控的视频水印方案. 上述视频 Hashing 算法的一个共同缺点是不能很好地抵抗针对视频的各种几何攻击.

当前针对几何攻击而设计的 Hashing 算法, 大多见于图像领域. 文献 [8] 直接利用 DCT 系数来构造 Hashing 函数, 对图像尺度变换、剪切等操作具有鲁棒性, 但对图像旋转不鲁棒. 文献 [9] 对图像实施 Fourier-Mellin 变换来计算 Hashing 值, 对仿射变换鲁棒. 文献 [10] 提出了基于 Zernike 矩的图像 Hashing 算法, 对图像旋转有一定的鲁棒性. 文献 [11] 利用图像傅里叶幅度谱极坐标变换来设计 Hashing 方案, 对图像的平移、拉伸、旋转有鲁棒性. 文献 [12] 则选择灰度直方图形状作为图像的特征, 取得了很好的抗几何失真性能. 虽然图像 Hashing 函数可以通过逐帧计算它们的 Hashing 值并串接起来的方法来扩展到视频领域, 但由于这种方法忽略了视频数据的时间关联性, 由此构造的视频 Hashing 算法往往会对帧率变化、帧丢失等时域同步攻击十分敏感, 而且计算复杂度很高, 难以满足实时性要求. 可见, 抗几何失真的视频 Hashing 算法要取得更好性能, 仍有待进一步深入研究.

本文提出一种基于灰度直方图特征的视频 Hashing 方案, 其鲁棒性原理在于: 1) 通过对视频数据实施时间轴方向的滤波处理, 使帧间变化信息扩散, 从而减轻了帧率变化、帧丢失等时域同步攻击对 Hashing 值的影响; 2) 对视频帧的滤波操作有效地增强了算法对噪声干扰、模糊滤波、有损压缩等处理的鲁棒性; 3) 利用直方图形状对几何失真的稳定性, 保证了算法对几何失真的强鲁棒性. 同时, 由于采用的特征是视频的灰度直方图, 算法因此具有较低的计算复杂度. 实验结果表明, 所提出的 Hashing 算法在抗各种几何攻击和视频处理操作上提供了良好的鲁棒性能.

本文第 2 节介绍所提算法的具体流程, 第 3 节详细阐述算法的鲁棒性原理, 第 4 节是算法的性能测试数据, 第 5 节是结论.

2 Hashing 方案

本文提出的基于视频灰度统计特征的 Hashing 算法的流程如图 1 所示.

Hashing 值的计算步骤可表述如下:

- (1) 将输入的视频数据解码后, 转换成灰度帧集 $V(w, h, f)$, w, h 表示视频的帧宽和帧高 (以 pixel 计), f 表示视频的帧数.
- (2) 预处理过程包括时域滤波抽样和空域滤波.

图 1 本文的 Hashing 算法流程

Figure 1 Procedure of the proposed algorithm LHVH

图 2 时域滤波抽样过程

Figure 2 Procedure of temporal filtering followed by sampling

时域滤波抽样称每一帧相同位置的像素构成的序列为像素通道 (pixel tube). 坐标为 (m, n) 的像素构成的像素通道表示为

$$PT(m, n) = \{V(m, n, i) | i = 1, 2, \dots, f\}. \quad (1)$$

对每一个像素通道进行方差为 σ_1^2 的一维高斯低通滤波 (此过程称为时域滤波), 再从时域滤波后的帧集中等间隔抽取 F 帧 (此过程称为时域抽样), 得到帧数归一化后的灰度帧集 $V(w, h, F)$ (实施过程如图 2 所示).

空域滤波对 $V(w, h, F)$ 的每一帧进行方差为 σ_2^2 的二维高斯低通滤波, 得到灰度帧集 $V'(w, h, F)$.

(3) 计算帧集 $V'(w, h, F)$ 的灰度均值 $\bar{A}$ 和灰度直方图 $\{h(i) | i = 0, \dots, 255\}$, $h(i)$ 表示灰度为 i 的像素个数, 并按如下方式截取直方图作为视频特征:

$$\{h(i) | i \in [\alpha, \beta], \alpha = \lfloor (1 - \lambda) \cdot \bar{A} \rfloor, \beta = \lfloor (1 + \lambda) \cdot \bar{A} \rfloor\}, \quad (2)$$

(2) 式中 $\lfloor X \rfloor$ 表示不大于 X 的最大整数. 实验发现, 直方图两端的灰度级像素数目易受几何失真影响, 因此, 参照灰度均值 $\bar{A}$ 提取直方图可避免将直方图两侧灰度级的统计数据用作视频特征, 有利于提高算法的鲁棒性. 经实验测试, 取 $\lambda \in [0.4, 0.6]$, 算法可获得较佳的性能.

(4) 在 $[\alpha, \beta]$ 中等间隔划分出 L 个子区间, 每个子区间含有灰度级的个数为

$$M = \left\lfloor \frac{\beta - \alpha + 1}{L} \right\rfloor. \quad (3)$$

将落在同一子区间的像素个数累加, 得到一个 L 维的特征矢量:

$$\left\{ h_M(i) | h_M(i) = \sum_{j=M(i-1)}^{M \cdot i - 1} h(j + \alpha), i = 1, \dots, L \right\}. \quad (4)$$

比较特征矢量不同元素两两间的大小关系来生成 Hashing 比特:

$$bit = \begin{cases} 1, & \text{if } h_M(i) \geq h_M(j), \\ 0, & \text{otherwise,} \end{cases} \quad 1 \leq i \leq j \leq L. \quad (5)$$

Hashing 比特串的长度为 $C_L^2 = \frac{L(L-1)}{2}$. 在密钥 K 的控制下对该比特串实施置乱操作, 最终得到视频的 Hashing 值.

3 鲁棒性原理

本文提出的视频 Hashing 算法的鲁棒性原理在于: 通过使用时域和空域高斯低通滤波器, 保证了所提算法对视频时间轴操作和帧内滤波操作的鲁棒性; 而以直方图形状作为视频特征, 则保证了所提算法对几何攻击的鲁棒性.

3.1 高斯低通滤波器的使用

文献 [6,11~13] 的研究亦表明, 高斯低通滤波是增强 Hashing 算法对模糊滤波、有损压缩等常见内容保持操作鲁棒性的一种普遍适用的手段. 为使得本文所提 Hashing 算法对噪声干扰、有损压缩等具有鲁棒性, 在算法的预处理步骤中进行了时域和空域上的高斯低通滤波:

I. 时域过滤: 对每一个像素通道实施时间轴方向上一维的高斯低通滤波, 可以平滑时间轴上亮度的微小变化, 同时将帧间急剧的变化与物体的运动特征扩散到多个帧中, 有利于提高 Hashing 算法对视频时间轴操作的鲁棒性.

II. 空域过滤: 对经时域滤波抽样处理后的每一帧进行二维的高斯低通滤波, 滤除了帧内不代表帧感知内容但易发生变化的高频成分, 从而有效增强了算法对视频模糊滤波等常用帧内处理的鲁棒性.

高斯滤波过程数学上可表述为高斯核函数 G 与灰度数据 I 的卷积和运算. 以二维高斯滤波为例, 可表述如下:

$$F(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \quad (6)$$

式中 $F(x, y, \sigma)$ 表示滤波后的灰度数据. 对 G 进行离散化时, 考虑到高斯核函数在距离其均值 3σ 处时函数值已下降至峰值的 1.1%, 因此, 选取高斯掩模的大小为 $(2 \times \lceil 3\sigma \rceil + 1) \times (2 \times \lceil 3\sigma \rceil + 1)$ 则有足够的运算精度 ($\lceil X \rceil$ 表示不小于 X 的最小整数).

选取合适的 σ 实施高斯低通滤波是能否有效增强算法性能的关键. σ 太大, 将滤除视频大量的动态内容, 会导致唯一性的下降; σ 过小, 视频内容的高频成分未能被有效抑制, 会影响鲁棒性的提高. 经实验观察, 选择 $\sigma_1 = 2$ 实施帧间滤波, $\sigma_2 = 3$ 进行帧内滤波, Hashing 算法具有较佳的性能.

3.2 稳健的直方图特征

从机理上看, 几何攻击是对像素位置的变换操作, 并不改变像素的颜色值^[14]. 视频的全局灰度直方图与像素的位置信息无关, 因此, 以直方图形状为特征来计算 Hashing, 可获得对视频几何失真的强鲁棒性. 由于提取直方图特征时是参照灰度均值进行的, 因此先讨论几何攻击下视频灰度均值的稳定性.

图 3 (a) foreman 原视频及其受攻击版本; (b) 循环平移 20%; (c) 尺度缩小至原来的 80%; (d) Loose 模式旋转; (e) 帧内裁剪 20%; (f) Crop 模式旋转; (g) 扭曲

Figure 3 (a) Original version; (b) circular shifting 20%; (c) scaling 80%; (d) loose rotation 20°; (e) intra-frame cropping 20%; (f) crop rotation 20°; (g) shearing 10%

表 1 受几何攻击后, 灰度均值的相对变化情况

Table 1 Relative change of mean under six geometric distortions

Operation type	Circular shifting 20%	Scaling 80%	Loose rotation 20°	Cropping 20%	Crop rotation 20°	Shearing 10%
Relative change	0%	0.1%	-5.23%	-1.81%	-5.43%	-5.30%

3.2.1 灰度均值的稳定性

对标准测试视频 foreman 实施循环平移 (circular shift)、尺度拉伸 (scaling)、Loose 模式旋转 (rotation)、帧内裁剪 (cropping)、Crop 模式旋转 (rotation+cropping)、扭曲 (shearing) 共 6 种常见的几何攻击 (攻击效果如图 3 所示). 受攻击前后视频帧集灰度均值的相对变化情况如表 1 所示.

表 1 的数据显示, 对标准视频 foreman 进行大角度旋转、裁剪等较强的几何攻击后, 视频帧集的灰度均值相对于原视频的灰度均值, 最大偏离只 5% 左右, 表明在受到几何攻击时, 视频帧集的灰度均值有良好的稳定性. 这是参照灰度均值提取直方图作为视频特征的前提.

3.2.2 直方图形状对几何失真的稳定性

设置算法参数为: $F = 64$, $\sigma_1 = 2$, $\sigma_2 = 3$, $\lambda = 0.6$, $L = 30$, 对视频样本 foreman 进行如图 3 所示的 6 种几何攻击. 然后从得到的特征矢量 $\{h_M(i)|i = 1, \dots, 30\}$ 中随机选择 30 对两两不同的元素作商, 所得结果如图 4 所示.

这些元素比值的稳定性反映了直方图形状的稳定性. 对视频实施帧循环平移和尺度拉伸处理, 所得的特征矢量元素的比值与原视频特征矢量元素的比值非常吻合, 表明在这两种处理下直方图形状十分稳定. 在其余 4 种几何攻击下, 特征矢量元素的比值在允许的范围内有微小的扰动. 由此可见, 以直方图形状为特征的感知 Hashing 算法能对几何失真提供很好的鲁棒性能.

4 实验结果及分析

4.1 唯一性测试

设置算法参数为 $F = 64$, $\sigma_1 = 2$, $\sigma_2 = 3$, $\lambda = 0.6$, $L = 30$. 计算包括 container, foreman, salesman

图 4 6 种攻击下, 对特征矢量随机选取的 30 对元素比值的变化情况 (+ 表示原视频特征矢量元素的比值, o 表示受攻击后特征矢量元素的比值)

Figure 4 Relative mean changes of thirty randomly selected element pairs under six distortions ('+' stand for element pairs of original video, 'o' stand for element pairs of distorted video)

等标准测试视频在内的共 200 段视频的 Hashing 值 (这些视频中有 20 段标准测试视频, 其余 180 段来源于网络, 包括自然风光、人物访谈、科普教育等内容, 持续时间由 6 s 到 25 s 不等, 分辨率为 QCIF, CIF 两种). 归一化汉明距是一种衡量 Hashing 算法唯一性的技术指标^[11], 其计算公式如下:

$$D(\text{hash}_1, \text{hash}_2) = \frac{1}{C_L^2} \sum_{k=1}^{C_L^2} |\text{hash}_1(k) - \text{hash}_2(k)|. \quad (7)$$

式中, hash_1 和 hash_2 表示两个不同的 Hashing 值, $\text{hash}_1(k)$ 和 $\text{hash}_2(k)$ 代表 hash_1 和 hash_2 的第 k 个比特. 对于两段不同的视频, 汉明距应在 0.5 附近; 对于两段内容上相似的视频, 其值应接近于 0. 测试所用 200 段视频数据两两间的归一化汉明距的统计直方图如图 5 所示.

图 5 所示数据的均值为 0.4530 (两个随机构造的 Hashing 值的汉明距期望是 0.5)、方差为 0.0395, 汉明距小于 0.15 的比例低于 0.1, 表明所提的 Hashing 算法具有良好的唯一性. 注意到仍有部分数据落在小于 0.15 的区间, 主要原因是本文算法使用的直方图特征是视频的一种低层特征, 尚未能与视频的语义有效关联, 因此整体灰度相似的视频的 Hashing 值可能发生碰撞. 与认证 Hashing 算法的强无碰撞性要求不同, 检索 Hashing 算法的任务在于根据视频内容筛选出候选视频, 允许存在较低的碰撞率.

4.2 鲁棒性测试

设置算法参数为: $F = 64$, $\sigma_1 = 2$, $\sigma_2 = 3$, $\lambda = 0.6$, $L = 30$. 对 20 段标准测试视频进行多种帧间和帧内处理, 得到的鲁棒性数据如表 2 所示 (通常情况下, 攻击前后 Hashing 值的归一化汉明距若小于

图 5 200 段视频 Hashing 值两两间的归一化汉明距的统计直方图

Figure 5 Distribution of NHDs (totally 19900 samples derived from 200 video clips)

0.15, 则认为算法对这种攻击是稳健的).

在表 2 中, 我们列出了 14 种视频操作前后 Hashing 值的变化情况, 对测试结果可作如下分析:

1) 对高斯白噪声干扰、高斯低通滤波处理具有较好的鲁棒性. 随着噪声强度和滤波强度的增加, 归一化汉明距逐渐增大. 在标准差为 15 的高斯噪声干扰和标准差为 5 的高斯低通滤波操作下, 平均汉明距小于 13%.

2) 能抵抗一定程度的亮度变化操作. 在 $\pm 5\%$ 亮度操作 (全局亮度提升和下降) 下, 汉明距值在 13% 以内. 亮度变化越大, 对汉明距的影响越明显, 主要原因是亮度变化直接导致灰度均值发生扰动, 使得攻击前后根据灰度均值截取的直方图形状特征可能不一致.

3) 对时间轴裁剪、帧丢失、帧率变化等时间轴操作十分稳健. 即使在大幅度的时间轴操作攻击下, 归一化汉明距也小于 2%, 表明基于直方图形状的视频 Hashing 对这些时间同步攻击的强鲁棒性.

4) 可以很好地抵御大角度旋转、高百分比扭曲和裁剪操作. 可以看出, 在 30° 的旋转下 (包括 Crop 和 Loose 两种模式)、裁剪去 20% 的内容和 20% 的扭曲下, 其汉明距值仍然在 15% 以内, 表明 Hashing 算法对这些几何攻击十分鲁棒.

5) 对帧内循环平移和 MPEG4 有损压缩有强鲁棒性. 帧内循环平移主要通过对像素在视频中位置进行移位引起失真, 而 MPEG4 有损压缩主要影响视频内容的高频成分. 由于直方图与像素位置无关且从低频成分中提取, 因而对这两种攻击鲁棒.

此外, 常见的视频处理操作还有视频内容编辑、对比度改变等. 视频内容编辑通过剪切物体, 添加效果等方式对视频内容进行修改, 会对视频全局灰度直方图形状造成一定的扰动. 由于全局直方图是整个视频片段的统计特性, 因此, 如果编辑区域不大, 或是编辑后的区域与原区域直方图相似, 则本文算法仍能生成相近的 Hashing 值. 对比度改变操作是对视频亮度的直接修改操作, 因此可能会极大地改变直方图形状, 本文算法对此不具有鲁棒性.

5 结论

本文提出一种基于灰度直方图抗几何失真的鲁棒视频 Hashing 方案, 主要工作总结如下:

表 2 鲁棒性测试数据

Table 2 Robustness testing results

Operation types	Param.	Container	Foreman	News	Salesman	Mean (20 clips)
Gaussian noise disturbance (noise deviation)	5	0.0069	0.0575	0.0161	0.0046	0.0280
	10	0.0299	0.1839	0.0391	0.0138	0.0679
	15	0.0414	0.2000	0.0828	0.0299	0.1224
Gaussian low-pass filtering (kernel deviation) ^{a)}	1	0.0115	0.0138	0.0322	0.0069	0.0144
	3	0.0552	0.1057	0.0782	0.0460	0.0622
	5	0.0506	0.2207	0.1356	0.1057	0.1107
Global luminance increase (increase percent)	1%	0	0	0	0	0.0147
	3%	0.0276	0.0874	0.0391	0.0414	0.0683
	5%	0.0667	0.1287	0.1287	0.1195	0.1233
Global luminance decrease (decrease percent)	1%	0.0161	0.0299	0	0	0.0114
	3%	0.0483	0.1632	0.0782	0.0391	0.0785
	5%	0.0621	0.2069	0.1126	0.0759	0.1221
Time-axis cutting (percent cut in two ends)	10%	0.0046	0.0276	0.0161	0.0023	0.0154
	15%	0.0069	0.0299	0.0207	0.0069	0.0202
Frame random dropping and restoring (dropping percent)	50%	0.0023	0.0115	0.0092	0.0115	0.0060
	70%	0.0046	0.0437	0.0184	0.0069	0.0114
Frame rate change (change ratio)	1/2	0.0023	0.0391	0.0161	0.0023	0.0199
	2	0.0023	0.0115	0.0138	0.0069	0.0100
Frame scaling (scaling ratio)	1/2	0.0414	0.0874	0.1885	0.1333	0.1179
	2	0.0414	0.0621	0.1011	0.1057	0.0687
Frame rotation (Loose) (rotation degree)	20°	0.0345	0.0897	0.0460	0.0460	0.1032
	30°	0.0391	0.0897	0.0483	0.0437	0.1010
Frame rotation (Crop) (rotation degree)	20°	0.0621	0.1080	0.0759	0.0230	0.1003
	30°	0.0621	0.1172	0.0805	0.0575	0.0908
Intra-frame cropping (cropping percent)	10%	0.0138	0.0368	0.1494	0.0207	0.0459
	20%	0.0391	0.0759	0.2000	0.0644	0.1118
Frame shearing (shearing percent)	10%	0.0414	0.0966	0.0391	0.0460	0.1040
	20%	0.0345	0.1103	0.0414	0.0736	0.1070
Circular shifting (shift percent)	30%	0	0	0	0	0
	50%	0	0	0	0	0
MPEG4 lossy compression (compression quality) ^{b)}	medium	0.0230	0.0391	0.0736	0.0138	0.0380
	low	0.0253	0.0391	0.0575	0.0161	0.0354

a) (参数: 高斯核函数标准差); b) (参数: 压缩质量)

1) 根据全局灰度直方图与像素位置无关的原理, 通过实验分析了直方图形状在常见几何攻击下的稳定性. 实验表明, 参照灰度均值提取到的直方图形状在几何攻击下十分稳定, 可作为鲁棒特征用于

视频 Hashing.

2) 利用直方图形状对几何失真的稳定性, 并结合时空维的高斯低通滤波来设计 Hashing 算法, 在对几何攻击鲁棒的同时, 也保证了对常见的视频帧内和帧间操作的鲁棒性.

3) 全局灰度直方图易于计算. 因此, 所设计的算法实现简单, 具有较低的计算复杂度.

本文所提的 Hashing 算法具有良好的唯一性和鲁棒性, 但由于只使用了灰度的全局统计特性, 因此对某些内容差异较大但灰度直方图相似的视频来说, 区分性不大. 不同于单幅图像, 视频通常由多个场景构成. 联合多个场景进行检索, 可以进一步降低本文算法的误检率. 当用于视频内容认证时, 安全性是另一个需要考虑的重要指标. 本文通过对 Hashing 比特进行随机置乱获得了一定的安全性, 但要满足视频认证的安全性要求, 还有待对直方图特征进一步分析和改善, 这将会是下一阶段的工作.

参考文献

- 1 Hamon K, Schmucker M, Zhou X. Histogram-based perceptual hashing for minimally changing video sequences. In: Proceedings of the 2nd International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution. Leeds, 2006. 236–241
- 2 Harris C, Stephens M J. A combined corner and edge detector. In: Proceedings of the 4th Alvey Vision Conference. Manchester: University of Manchester, 1988. 147–152
- 3 Zhou Y, Liu G, Dai Y, et al. Robust hashing based on persistent points for video copy detection. In: 2008 International Conference on Computational Intelligence and Security. Suzhou, 2008. 305–308
- 4 Zhou X, Schmucker M, Brown C L. Perceptual hashing of video content based on differential block similarity. LNCS, 2005, 3802: 80–85
- 5 Lee S, Yoo C D. Robust video fingerprinting for content-based video identification. IEEE Trans Circ Syst Video Tech, 2008, 18: 983–988
- 6 Coskun B, Sankur B, Memon N. Spatio-temporal transform based video hashing. IEEE Trans Multimedia, 2006, 8: 1190–1208
- 7 Liu L, Peng D, Li X. A security video watermarking scheme for broadcast monitoring. In: Proceedings of IWSDA'07. Chengdu, 2007. 109–113
- 8 Zhang W K, Kong X W, You X G. Secure and robust perceptual hashing. J South-East Univ (Nat Sci Ed), 2007, 37: 188–192 [张维克, 孔祥维, 尤新刚. 安全鲁棒的图像感知 Hashing 技术. 东南大学学报 (自然科学版), 2007, 37: 188–192]
- 9 Swaminathan A, Mao Y, Wu M. Image hashing resilient to geometric and filtering operations. In: IEEE Workshop on Multimedia Signal Processing. Siena, 2004
- 10 He M Y. Study of robust and secure digital image signature. Master Thesis. Guangzhou: Sun Yat-sen University, 2007 [何妙谊. 稳健及安全的数字图像签名研究. 硕士学位论文. 广州: 中山大学, 2007]
- 11 Swaminathan A, Mao Y, Wu M. Robust and secure image hashing. IEEE Trans Inf Forens Secur, 2006, 1: 215–230
- 12 Xiang S J, Kim H J, Huang J W. Histogram-based image hashing scheme robust against geometric deformations. In: Proceeding of the 9th ACM Multimedia and Security Workshop. Dallas, 2007. 121–128
- 13 Qin C, Wang S Z, Zhang X P. Image hashing based on human visual system. J Image Graph, 2006, 11: 1678–1681 [秦川, 王朔中, 张新鹏. 一种基于视觉特性的图像摘要算法. 中国图象图形学报, 2006, 11: 1678–1681]
- 14 Xiang S J, Kim H J, Huang J W. Invariant image watermarking based on statistical features in the low-frequency domain. IEEE Trans Circ Syst Video Tech, 2008, 18: 777–790

Perceptual video Hashing robust against geometric distortions

XIANG ShiJun^{1,4}, YANG JianQuan² & HUANG JiWu^{3*}

¹ School of Information Science and Technology, Jinan University, Guangzhou 510632, China;

² Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China;

³ School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China;

⁴ State Key Laboratory of Information Security (Institute of Information Engineering, Chinese Academy of Sciences), Beijing 100029, China

*E-mail: isshjw@mail.sysu.edu.cn

Abstract In this paper, we propose a robust perceptual hashing algorithm by using the shape of video luminance histogram. The underlying robustness principles are based on three main aspects: 1) since the histogram in shape does not depend on the exact position of a pixel, the algorithm is resistant to geometric deformations; 2) the hash is extracted from the spatial Gaussian-filtering low-frequency component against common video processing operations such as noise permutation, low-pass filtering and lossy compression; 3) temporal Gaussian-filtering operation is designed so that the hash is robust to temporal desynchronization operations, such as frame rate change and dropping. As a result, the hash function is robust to common geometric distortions and video processing operations. Experimental results show that the proposed hashing strategy provides satisfactory robustness and uniqueness.

Keywords video Hashing, geometric distortion, histogram, statistical features, Gaussian filtering

XIANG ShiJun received his B.S. degree from Chang'an University (1997), M.S. degree from Guizhou University (2000) and Ph.D. degree from Sun Yat-Sen University (2006). Currently he is an Associate Professor with Jinan University, Guangzhou. His current research interests include digital watermarking, multimedia hashing and other security issues in digital media.

YANG JianQuan received the B.S. and M.S. degrees from Sun Yat-sen University in 2008 and 2010, respectively. He is currently a Research Assistant at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His current research interests include multimedia security and image processing.

HUANG JiWu received the B.S. degree from Xidian University, Xi'an in 1982, the M.S. degree from Tsinghua University, Beijing in 1987, and Ph.D. degree from Institute of Automation, Chinese Academy of Sciences, Beijing in 1998. He is currently a Professor with the School of Information Science and Technology, Sun Yat-Sen University, Guangzhou. His current research interests include multimedia security and data hiding. Dr. Huang serves as an associate editor of IEEE Transactions on Information Forensics and Security, a member of the IEEE CASS Technical Committee of Multimedia Systems and Applications, a member of IEEE SPS Technical Committee of Information Forensics and Security.