

基于云水印的视频内容认证技术*

梁长垠¹, 李昂², 牛夏牧²

(1. 深圳职业技术学院 电子与信息工程学院, 广东 深圳 518055;
2. 哈尔滨工业大学 深圳研究生院, 广东 深圳 518055)

摘要: 数字视频内容的认证技术越来越成为研究的热点。基于云水印及镜头分割技术, 提出了一种新颖的视频内容认证方案。论文就不变特征提取、云水印生成、嵌入和提取、视频的认证与篡改检测等技术进行了研究并进行了实现。通过实验分析和验证, 该方案可对实时的视频进行认证而不影响视频本身的视觉效果。

关键词: 视频内容认证; 云水印; 镜头分割; 特征提取

中图分类号: TP309.2 **文献标识码:** A **文章编号:** 0529-6579(2009)01-0026-05

Video Content Authentication Technique Based on Cloud Watermark

LIANG Changyin¹, LI Ang², NIU Xiamu²

(1. School of Electronics & information Engineering, Shenzhen Polytechnic, Shenzhen 518055, China;
2. Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China)

Abstract: Digital video content authentication technique becomes a hot point. A novel video content authentication scheme, based on cloud watermark and shot segmentation, is proposed. The related technique on the Invariant Feature extraction, the generation of the cloud watermark, the embedding and the extraction of cloud watermark, the authentication and the tamper detection on video were studied and implemented. With some experiments and the corresponding analysis and the verifying for the results, the scheme can be applied into the real time video and may not affected the visual efficient.

Key words: video authentication; cloud watermark; shot segmentation; feature detection

随着近年来网络多媒体技术的飞速发展, 数字视频在网络中传播的越来越广泛, 相应地, 视频数据遭受恶意篡改所造成的社会危害也越来越大。基于此, 作为保证视频内容完整性和真实性的视频内容认证技术成为近来研究和应用的热点。而半脆弱性数字水印技术因其具有隐蔽性好、对失真具有一定容忍度、不需要额外保存认证信息、具有篡改定位能力等优良特性, 在多媒体信息的内容认证方面具有很好的应用前景。

目前, 文献[1]结合LSB脆弱水印算法和基于扩频的视频鲁棒水印算法, 设计了针对MPEG视频认证的半脆弱水印系统, 该水印系统能够抵抗码率转换等内容保持类操作, 但水印生成时计算复

杂, 实时性差。文献[2]提出了一种自恢复认证的云水印算法, 该算法利用修改VLC(Variable length coding)参数在压缩视频码流中嵌入水印, 将后一个图像组(GOP_{i+1})的主要特征值压缩后嵌入当前图像组(GOP_i)中。这虽然能够对图像的一般修改进行认证与恢复, 但不能抵制转码、加噪等内容保持类的操作。文献[3]提出了一种基于内容特征的可配置视频认证方案, 该方案虽然可以通过参数调节的方式实现系统鲁棒性与安全性之间的较好平衡, 但同样对改变视频结构的内容保持类操作敏感。本文将结合视频内容的特点, 应用镜头分割技术对视频进行不变特征提取, 提出一种基于云水印的视频水印认证方案^[4-5]。

* 收稿日期: 2008-02-25

基金项目: 广东省自然科学基金资助项目(05109511); 广东省科技计划资助项目(2006B37430001)

作者简介: 梁长垠(1964年生), 男, 教授; E-mail: szcylang@hotmail.com

1 水印算法设计

为了区别恶意篡改与正常的视频处理过程, 本文设计了一种基于云水印的视频认证技术。该技术基本思想是: 压缩的视频码流首先通过部分解码以得到 DCT 系数, 嵌入水印后再重新编码。然后通过接收视频中提取的水印信息与重构的特征信息进行比较, 以实现对时域及空域恶意篡改的检测。

1.1 不变特征提取

在图像认证方面, 已经有各种各样从图像中提取不变特征量的鲁棒哈希算法被提出^[6-7]。而将这些算法扩展到视频认证中时却存在着一些问题。由于视频通常是由大量的图像帧构成, 复杂的图像特征提取算法会导致水印嵌入耗时过长, 不能满足视频认证的实时性要求。并且这些视频帧的特征量之间也缺乏联系, 对时域上的篡改检测无能为力。考虑到在绝大多数的视频序列中, 一个镜头内连续的图像帧在视觉上都是十分相似的特点, 本文将待认证的视频进行镜头分割^[8-9], 以镜头为单位提取不变特征量。镜头分割方法的主要差别在于帧与帧之间的比较方法不同。首先, 从视频帧中提取某种特征来衡量帧间内容的变化, 例如颜色、形状、纹理等特征都可以用来作为对视频帧的描述; 然后, 根据某种标准来度量帧间内容的差别, 这种差别就是计算出来的帧间不连续值 $Z(k, k+1)$, 其中 $k \geq 1$, 最后经过阈值判决, 不连续值大于某一预先设定的阈值, 则认为帧 k 与 $k+1$ 形成了边界, 否则两帧在同一镜头内。

对 MPEG-2 压缩的视频来说, 一方面由于中、低频 DCT 系数集中了图像帧的绝大部分能量, 较好地反映了视频的内容变化; 另一方面, 中、低频系数在重压缩等合法操作的过程中相对稳定, 而对视频帧内容上的篡改敏感^[10]。因此利用视频帧亮度块内中、低频 DCT 系数的能量分布作为镜头分割的依据及内容特征, 是一个便捷且恰当的选择。

设一个视频帧中有 m 个宏块, 每个宏块包含 4 个亮度块, 每个亮度块都由 64 个 DCT 系数组成。基于 DCT 系数能量的镜头分割方法可分为以下三个步骤:

(1) 首先对视频序列内每帧图像的所有宏块分别计算其亮度块的中低频 DCT 系数能量和 EM , 即 $EM = \sum_{i=0}^3 \sum_{j=9}^c \theta_j^2$ 。其中, θ_j 表示块中的 DCT 系数, j 表示 DCT 系数经过 Zig-Zag 扫描后的系数序号, 系数序号小于 c 的 DCT 系数称为中低频系数。

(2) 计算第 k 帧图像与第 $k+1$ 帧之间的特征

距离 $d(k, k+1) = \sum_{i=1}^m |EM_{i,k+1} - EM_{i,k}|$, 这个距离用来度量两个连续帧之间的相异程度。

(3) 将距离 $d(k, k+1)$ 与一个阈值 T_s 进行比较, 如果 $d(k, k+1) > T_s$, 则第 k 帧为一个镜头的边界, 阈值 T_s 的选取可参考文献 [11]。

对视频序列进行镜头分割后, 将一个镜头 S 内所有 I 帧的中低频 DCT 能量块做平均, 产生一个有代表性的能量分布描述矩阵 X , r 为镜头 S 内的 I 帧总数, 其中, $X_i = \frac{1}{r} \sum_{j=1}^r EM_{i,j}$ ($i = 1, 2, \dots, m$)。

接下来将矩阵 X 均匀划分为一系列子区域 $\{R_1, R_2 \dots R_n\}$, n 为子区域数量, 也即该镜头特征向量的长度。每个子区域包括 m/n 个宏块。计算序列 $\{R_1, R_2 \dots R_n\}$ 中每个子区域的能量值, 来产生一个新的序列 $\{E_1, E_2 \dots E_n\}$, 并获得这个序列元素的中间值 E_d , 其中 $E_i = \sum_{j \in R_i} EM_j$ ($i = 1, 2, \dots, n$)。

将这个能量序列归一化为一个二值形式, 得到镜头 S 的特征向量 F :

$$F(i) = \begin{cases} 0, & E_i < E_d \\ 1, & E_i \geq E_d \end{cases} \quad (i = 1, 2, \dots, n)$$

特征量 F 很好地反映了镜头 S 内所有 I 帧图像的能量分布情况。这里也可以引入其它一些复杂的图像特征提取技术, 以期进一步地提高认证水印的篡改检测能力。

1.2 云水印的生成

为增加安全性, 可用密钥 K 将特征量 F 置乱并进一步做适当的调制, 产生 m 位的水印信息 $W \{w_1, w_2, \dots, w_n\}$, 其中 w_i 是特征位 $F(i)$ 经扩频后得到的 m/n 位伪随机序列。根据云模型及视频序列的特性, 通过扩展云发生器^[4-5]生成嵌入到镜头 S 中的云水印信息。

以水印信息集合 $W \{w_1, w_2, \dots, w_n\}$ 为期望值, 选择适当的 En 和 He , 可以产生一系列相似但不同的水印云滴的集合 $\{D_1, D_2, \dots, D_t\}$, 其中 $D_i = W'$, t 为镜头 S 包含的视频帧总数。具体的实现过程如下:

(1) 以 En 为期望, He^2 为方差, 生成一个随机熵 En' ;

(2) 以集合 W 中的元素 w_i 为期望, En'^2 为方差, 生成一个随机量 w_i' ;

(3) 遍历 W 中的所有元素, 产生一个水印云

滴 $D_i = W' \{w'_1, w'_2, \dots, w'_n\}$;

(4) 重复步骤(1) ~ (3) 直至产生 t 个云滴。

1.3 水印的嵌入和提取

镜头 S 对应的 t 个水印云滴将被逐一嵌回到该镜头所有帧内, 每帧图像将嵌入一个云滴。水印的嵌入方法则是通过调整宏块内的能量分布, 将水印嵌入在 DCT 系数的亮度块中, 每个宏块嵌入一个水印信息 bit 位。其中设定的能量阈值 T_E 用来提高水印的鲁棒性, 实际应用中可以根据具体需要调整 T_E 的大小, T_E 大, 则鲁棒性好, 图像视觉效果会相对降低一些; T_E 小, 则鲁棒性就差一些。具体嵌入方案如下:

(1) 在一个宏块内分别计算 4 个亮度块 B_0, B_1, B_2, B_3 的能量, 即 $EB_i = \sum_{j=0}^{63} \theta_j^2 (i = 0, 1, 2, 3)$;

(2) 当嵌入的水印信息为 0 时, 将块 B_1, B_2 的 DCT 高频系数清零并与块 B_0, B_3 交换部分中频系数, 直至 $(EB_0 + EB_3) - (EB_1 + EB_2) \geq T_E$;

(3) 当嵌入的水印信息为 1 时, 将块 B_0, B_3 的 DCT 高频系数清零并与块 B_1, B_2 交换部分中频系数, 直至 $(EB_1 + EB_2) - (EB_0 + EB_3) \geq T_E$;

通过这个阈值的设定, 可以保证宏块内两部分的能量差值始终不小于 T_E , 确保了水印对于常规视频处理的鲁棒性。

水印的提取是嵌入的逆过程, 分别计算接收视频每个宏块内四个亮度块的能量。若 $EB_0 + EB_3 \geq EB_1 + EB_2$, 则提取水印位“0”, 若 $EB_0 + EB_3 < EB_1 + EB_2$, 则提取水印位“1”。

1.4 认证及篡改检测

通过重构视频的内容特征并与从视频中提取出的水印信息进行比较, 可以实现对接收视频的空域及时域恶意篡改进行认证和检测定位。

首先将接收视频以同样方法重新分割为一系列的镜头, 从镜头 S' 内提取出 t' 个水印云滴, 利用逆向云发生器得到水印信息 $W' \{w'_1, w'_2, \dots, w'_n\}$ [4-5]。对镜头 S' 的内容特征进行重构得到向量 F' , 以相同的密钥 K 及伪随机序列置乱和调制后生成 m 位重构特征信息 $Z \{z_1, z_2, \dots, z_n\}$ 。分别计算伪随机序列 w'_i 与 z_i 之间的相关性, 可实现定位精度为 m/n 个宏块的空域篡改检测。

对一个逆向云发生器 [4-5] 来说, 依据统计原理, 给定的样本点越多, 逆向云算法的误差越小。在水印云滴样本点有限的情况下, 无论采用什么算

法, 恢复的水印信息 W' 与原始水印信息 W 之间的误差都是不可避免的 [5]。以样本平均值作为期望 Ex 估计值的算法方差为: $DX = \frac{1}{n^2} \sum_{i=1}^n En_i^2 =$

$(En^2 + He^2)/n$, 可看出: 误差范围是由样本数 n , 正向云发生器参数 En 和 He 决定的。因此设定一个阈值 T , 并在水印云滴生成时自适应地选取 $En = \sqrt{\frac{n}{3}}T, He = \frac{\sqrt{n}}{3}T$, n 即为一个镜头内视频帧数量。则 $W' \{w'_1, w'_2, \dots, w'_n\}$ 中任何一个 w'_i 的原始误差不会超过 $2/3T$, 阈值 T 的确定应以不影响空域的相关检测为前提。

时域上的篡改包括帧的剪切、插入及位置互换等, 这些攻击都会导致镜头内有效水印云滴样本的减少, 使恢复的水印信息 W' 与原始水印信息 W 之间存在的误差增大。以重构特征信息 $Z \{z_1, z_2, \dots, z_n\}$ 代替 W , 并忽略已检测到空域篡改的图像区域, 对 $W' \{w'_1, w'_2, \dots, w'_n\}$ 及 $Z \{z_1, z_2, \dots, z_n\}$ 的其余 n' 部分计算其平均距离:

$$d = \frac{1}{n'} \sum |w'_i - z'_i|$$

若 $d > T$, 则可判断镜头 S 内遭到了时域上的篡改。

在镜头层次上的时域篡改检测是通过不同的镜头改变置乱密钥来实现的, 所有的密钥均由一个状态机密钥发生器产生 [12]。在不知道密钥 K 的情况下进行水印的嵌入是几乎不可能的, 而即使知道了当前镜头所用的密钥, 也不能为其它镜头的密钥破解提供任何信息, 因此这种密钥转换的机制也进一步确保了认证水印方案的安全性。

2 实验结果及分析

为验证上一节中算法的可行性和有效性, 下面采用 MPEG-2 标准视频测试序列 Table-Tennis 进行实验, 并从对常规处理的鲁棒性、篡改检测能力等方面进行测试及讨论, 测试视频的主要参数, 视频大小为 704×480 ; GOP 帧数为 12; I/P 帧间距为 3; 色差格式为 4:2:0; 帧率为 25 fps; 帧数为 679。

2.1 常规处理后水印检测结果

视频文件由于其数据量极大, 并且与图像相比有较大的冗余, 因此在存储和传输的过程中通常要对其进行一些例如压缩码率转换、比例缩放、图像滤波等。其中, 尤以不同码率的转换最为常见, 包括适当去除视频序列的 B 帧或 P 帧、丢弃高频 DCT 系数、重新量化、将压缩视频转化为非压缩比特流后重编码。因此, 视频认证水印对于常规处

理,必须具有一定的鲁棒性。本文对算法在视频压缩重编码、码率转换、帧率转换后的鲁棒性进行了测试,如表 1 所示为 25 帧/s 与 20 帧/s 时的误码率。从中可以看出,平均每帧的误码率一般不超过 0.7 bit。当对视频进行篡改时,某帧或某些帧的误码率会很大。所以很小的误码率不影响对视频认证的效果。

表 1 误码率
Table 1 The false bit rate bit · 帧⁻¹

帧率/(帧 · s ⁻¹)	码率/Mbps			
	4	1.5	1.2	1
25	0.32	0.48	0.75	0.52
20	0.45	0.62	0.60	0.51

2.2 时域篡改攻击认证结果

为了评价认证方案对时域篡改检测能力,本文

共设计了如表 2 所示的 3 组实验。从其中时域认证结果可以看出,视频图像空域的内容篡改对时域的检测影响很大,原因在于对图像内容的篡改会使可用于时域检测的水印信息部分减少,计算误差增大。这会降低检测的正确率,同时会产生虚警,虚警率随着单帧篡改程度的加深和篡改帧数量的提高而变大。总体来说,本文提出的视频认证方案对时域上的篡改均有较好的检测效果。

2.3 空域篡改攻击认证结果

本文算法是对视频内容进行保护,实验结果如图 1 所示。从实验中可以看出,本文的认证水印方案可以检测出对视频图像的篡改,包括替换视频中的某个对象和在视频中删除或添加某个对象。通过本文提出的算法进行检测,可以比较精确地定位篡改位置(如图 1(c)中认证图像的黑块部分所示)。

表 2 时域认证结果

Table 2 The results of temporal authentication

对视频篡改程度	篡改类型	检测正确率/%	虚警率/%
对视频图像内容做轻微篡改的同时,对视频中部分片断进行删除和插入	删除	100	0
	插入	100	0.2
对视频图像内容做严重篡改的同时,对视频中部分帧做了删除和插入	删除	100	0.25
	插入	98	0.33
	删除	93	0.47
对视频图像内容做严重篡改(修改区域达到 50% 以上)的同时,对部分帧做了删除、插入和交换	插入	95	1.10
	交换	90	0.50

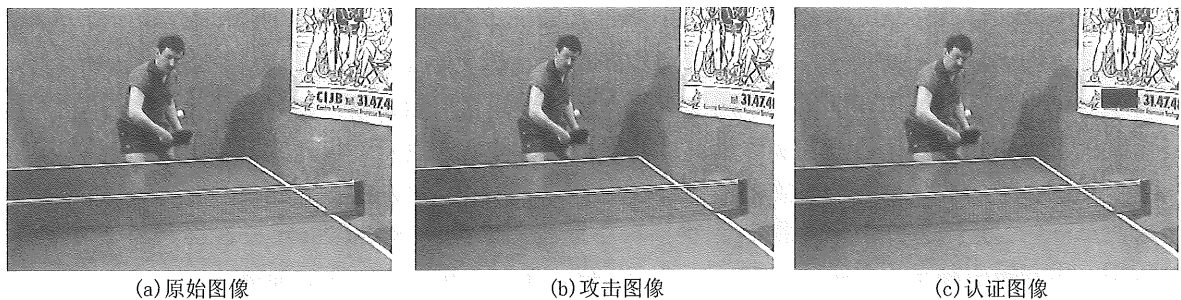


图 1 Table-Tennis 视频序列攻击后的检测结果

Fig. 1 The detection result after tempering for Table-Tennis

3 结 论

本文在视频镜头分割技术的基础上,提出了一种基于镜头的特征提取和云水印的视频内容认证方案,该方案可实现针对 MPEG 压缩视频进行内容篡改的检测及定位。通过对视频镜头进行特征提取,大大缩短了水印的嵌入时间,提高了认证算法

的实时性。利用有效水印云滴样本数量与还原误差之间的关系,检测视频内容在时域上遭受篡改的程度。基于能量分布的重构特征量与还原水印信息的相关性检测,能够对空域的篡改进行有效的定位。

参考文献:

- [1] PENG Yinyu, HEATHER Hong. A semi-fragile water-

- marking system for MPEG video authentication[J]. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, 2002, 4(5): 3461 - 3464.
- [2] CROSS Daniel, MOBASSERI Bijan. Watermarking for self-authentication of compressed video[C]. IEEE International Conference on Image Processing Rochester, New York, 2002, 3(9): 913 - 916.
- [3] SUN Q, HE D, TIAN Q A. Secure and robust authentication scheme for video transcoding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(10): 1232 - 1244.
- [4] 李德毅, 刘常昱, 杜鹃, 等. 不确定性人工智能[J]. 软件学报, 2004, 15(11): 1583 - 1594.
LI Deyi, LIU Changyu, DU Yi, et al. Artificial intelligence with uncertainty[J]. Journal of Software, 2004, 15(11): 1583 - 1594.
- [5] 李德毅, 杜鹃. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005.
- [6] YANG B, GU F, NIU X. Block mean value based image perceptual hashing[C]. International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2006: 167 - 172.
- [7] SWAMINATHAN A, MAO Y, WU M. Robust and secure image hashing[J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 215 - 230.
- [8] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003: 13 - 17.
- [9] 李国辉. 基于内容的多媒体信息检索模型和方法研究[D]. 长沙: 国防科学技术大学, 2001: 10 - 30.
- [10] MOBASSERI B G, EVANS A T. Content - dependent video authentication by self - watermarking in color space[C]. SPIE Security and Watermarking of Multimedia Contents III, San Jose, 2001, 4314(1): 21 - 26.
- [11] ROOVER C D, VLEESCHOUWER C D, LEFEBVRE F, et al. Robust video hashing based on radial projections of key frames[J]. IEEE Transactions on Signal Processing, 2005, 53(10): 4020 - 4037.
- [12] LIN E T, DELP E J. Temporal synchronization in video watermarking[J]. IEEE Transactions on Signal Processing, 2004, 52(10): 3007 - 3022.

· 简 讯 ·

本刊再次入选中国百种杰出学术期刊

据中国科技信息研究所 2008 年 12 月年度发布会公布的信息, 中山大学学报(自然科学版)再次被评为中国百种杰出学术期刊, 这是自 2002 年公布第一届中国百种杰出学术期刊名单以来, 本刊第五次获此殊荣。

中国百种杰出学术期刊是科技部直属单位中国科技信息研究所根据科技期刊综合评价的 19 项科学计量指标, 组织不同学科、不同层面的专家制订期刊指标体系, 确定重要指标的权重, 并根据我国学术期刊的变化趋势和实际状况, 分学科对每种期刊进行综合评定而推出的研究结果。

近年来, 中山大学学报(自然科学版)努力学习实践科学发展观, 狠抓学术质量的提高, 不断改革创新, 走跨越式发展之路, 刊物指标不断提高, 据中国期刊引证报告(2008 扩刊版)的数据, 本刊被引频次为 1641, 影响因子为 0.747, 比上一年度又有新的提高。

(本刊通讯员)