计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.201100089

基于宏块编码信息自适应置换的 H.264/AVC 视频加密方法

梁 剑 何军辉
华南理工大学计算机科学与工程学院 广州 510006
(doc.liang@qq.com)

摘 要 云存储的发展使人们愿意将个人视频数据传输至云端,但伴随而来的数据安全问题日益突出,选择加密是对视频进行隐私保护的有效手段之一。针对目前 H.264/AVC 视频选择加密方法普遍存在安全性不足的问题,文中提出了一种基于宏块编码信息自适应置换的 H.264/AVC 视频加密方法。该方法根据宏块的编码类型逐帧自适应生成伪随机序列,利用伪随机序列将宏块编码信息中的残差编码方案(Coded Block Pattern,CBP)和残差数据(Residual)在宏块间进行随机置换,同时还对 I 宏块的帧内预测模式以及 P 宏块与 B 宏块的运动向量差值的符号进行加密。实验结果表明,该方法可保证加密视频兼容 H.264/AVC 编码标准,并具有加密空间大、密钥敏感性好、视频码率变化小的特点。与现有的主流加密方案相比,所提方法在视觉安全性和抵抗最新提出的轮廓攻击方面表现更佳。

关键词:云服务;H.264/AVC;选择加密;编码信息;自适应置换
中图法分类号 TP309

H.264/AVC Video Encryption Based on Adaptive Permutation of Macroblock Coding Information

LIANG Jian and HE Jun-hui
School of Computer Science and Engineering,South China University of Technology,Guangzhou 510006,China

Abstract The development of cloud storage makes people willing to upload personal video to the cloud,but the data security problems brought by it have become increasingly prominent,selective encryption is one of the effective ways to protect video privacy. Aiming at the problem of insufficient security in the current H.264/AVC video selective encryption method,a novel H.264/AVC video selective encryption method based on adaptive permutation of macroblock coding information is proposed. The method adaptively generates pseudo-random sequence frame by frame according to the macroblock types,uses the pseudo-random sequence to randomly permute the coded block pattern (CBP) and the residual data in the coding information of a macroblock between macroblocks,changes the intra prediction modes of I macroblocks,and flips the signs of motion vector differences of P macroblocks and B macroblocks. Experimental results show that the proposed method can preserve format compatibility with H.264/AVC coding standard,and has characteristics of large encryption space,good key sensitivity,and small video bitrate variation. Compared with the existing encryption schemes,the proposed method performs better in terms of visual security and resistance to state of the art of sketch attack.

Keywords Cloud service,H.264/AVC,Selective encryption,Coding information,Adaptive permutation

1 引言

伴随着移动互联网的快速发展,直播、短视频、在线会议等各类视频应用井喷式爆发,越来越多的视频数据存储至云端。如何保护存储于公共云上的视频隐私成为人们关注的焦点,视频加密是解决云端视频安全的方法之一[1]。如图 1 所示,视频拥有者使用密钥对视频进行加密后上传至云端服务器,任何攻击者(包括不可信的云服务管理者)只能获取加密后的视频数据但无法查看视频内容,而视频拥有者可通过所拥有的密钥对从云端下载的加密视频进行正确解密,完全可逆地恢复原有视频数据,通过视频加密实现了视频数据在传输和存储过程中的隐私保护。

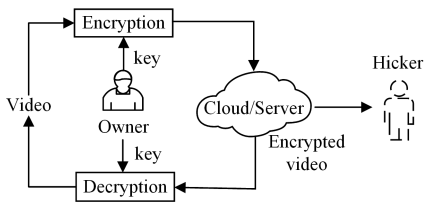


图 1 视频加密的应用场景

Fig.1 Application scenario of video encryption

到稿日期:2020-11-12 返修日期:2021-03-05 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:广东省自然科学基金(2019A1515011231)
This work was supported by the Natural Science Foundation of Guangdong Province(2019A1515011231).
通信作者:何军辉(hejh@scut.edu.cn)

由于未经编码的视频数据量极大，直接对其加密不但非常耗时，而且会严重破坏视频内容之间的相关性，使得编码器对加密后的视频数据进行压缩的效率较低。而对编码后的视频文件进行传统的数据加密会破坏格式兼容性，解码器无法对密文视频进行正常解码。视频选择加密只对码流中的关键数据进行加密，能够在保证与视频编码格式兼容的同时实现对数据隐私的保护，并且使得基于密文视频的信息隐藏、数字水印、内容检索等成为可能[2]。Ahn 等[3]根据 I 帧的重要性，提出了基于帧内预测模式（Intra Prediction Mode，IPM）的加密方法，但该方法无法针对 P 帧和 B 帧进行加密，加密效果不佳。Li 等[4]根据尺寸为 4×8 和 8×4 两种帧间预测子块所拥有的运动向量（Motion Vector，MV）个数相同的特点，随机置乱这两种预测子块的尺寸，加密元素非常有限，导致加密效果仍然不理想。Khlif 等[5]基于混沌加密算法来随机修改各个帧间预测子块的 MV 符号，在 P 帧和 B 帧上取得了较好的加密效果，但是 MV 符号的改变会导致运动向量差值（Motion Vector Difference，MVD）值的改变，进而影响其熵编码时的指数哥伦布码字字长，从而造成较为明显的视频码率增长。Su 等[6]利用符号相反且绝对值相同的两个数经过指数哥伦布编码后字长相同的特点，提出了对 MVD 进行符号随机置乱的加密方法，该方法在不影响码率的情况下获得了良好的加密效果。Shen 等[7]进一步将指数哥伦布码字按照字长进行分组，对 MVD 的码字进行组内随机置换，该方法相较于文献[6]不但能对 MVD 的符号进行干扰，还可能影响其绝对值。针对量化离散余弦变换（Quantized Discrete Cosine Transform，QDCT）系数的扫描顺序，Wang 等[8]提出了一种反 Zig-zag 扫描顺序，并在熵编码过程中针对 QDCT 系数随机选择 Zig-zag 或反 Zig-zag 顺序进行扫描。Ding 等[9]在文献[8]的基础上进一步分析 QDCT 系数中高频区域和低频区域对画面质量的影响，设计出了 8 种可供随机选择的 QDCT 系数扫描顺序，并通过实验证明新增的扫描顺序对编码压缩效率的影响较低。文献[10-16]不但对 IPM 和 MVD 进行加密，还对非零 QDCT 系数的符号进行随机置乱，能够在保证码率不变的同时产生更好的加密视觉效果，且相比 IPM 和 MVD，非零 QDCT 系数符号加密适用于所有帧类型。上述加密方法虽然能够在大多数情况下对视频内容进行保护，但易遭受最新提出的轮廓攻击[17]。

本文提出了一种新颖的 H.264/AVC 视频加密算法。首先逐帧根据宏块的编码类型自适应生成初始化向量（Initialization Vector，IV），并和用户密钥一起基于 AES-OFB 算法生成伪随机序列。然后将宏块编码信息分为两个部分：{宏块类型（mb_type），宏块预测（mb_pred），量化参数（Quantization Parameter，QP）}和{CBP，resi-dual}，并为一帧中除了分割为 16×16 的帧内预测类型（Intra_16×16）、跳跃类型（P_Skip 和 B_Skip）、直接类型（B_Direct）以外所有宏块的{CBP，residual}部分分配一个伪随机数。最后根据伪随机数对这些宏块的{CBP，residual}部分在宏块间进行随机置换加密。该方法相比现有的加密方案在视觉安全性和抵抗最新提出的轮廓攻击[17]方面得到了进一步的提升，同时具有与编码格式兼容、对密钥敏感、对视频码率影响小等特点。

本文第 2 节介绍了两种主流 H.264/AVC 视频加密方法，其部分思想同样被采纳以提高算法的安全性；第 3 节将具体描述本文所提出的置换加密方法；第 4 节从多个角度对加密算法进行性能分析；最后总结全文并展望未来。

2　相关工作

针对视频的视觉结构信息，Ahn 等[3]分析了分割为 4×4 的帧内预测类型（Intra_4×4）宏块的 IPM 的编码特点，提出了针对 IPM 的加密方法。在 H.264/AVC 编码标准中，Intra_4×4 类型宏块被进一步划分为 16 个尺寸为 4×4 的预测子块，且各预测子块有 9 种帧内预测模式可选[18]。若当前预测子块的帧内预测模式等于其上侧和左侧相邻预测子块的最小值，则只需使用 1 bit 字段 prev_intra 4×4_pred_mode_flag 记录当前子块的 IPM，否则需要额外的 3 bit 字段 rem_intra 4×4_pred_mode 表示。因此，从伪随机序列中读取 3 bit 数据，通过式(1)对宏块的 IPM 随机加密。

$$Mode_{\text{new}}=Mode_{\text{org}}\oplus R \tag{1}$$

其中，⊕为异或操作，$Mode_{\text{org}}$ 和 $Mode_{\text{new}}$ 分别表示加密前后宏块的 rem_intra4×4_pred_mode 字段，R 为 3bit 随机数据。该方法能够有效地破坏视频帧的图像视觉结构，且由于异或前后 rem_intra4×4_pred_mode 字段长度不变，所以不会带来任何视频码率的影响。

针对视频的运动信息，Su 等[6]提出了一种基于 MVD 符号置乱的加密方法。为了进一步降低视频运动信息之间的相关性，H.264/AVC 编码器首先对每个帧间预测子块计算其最优 MV，然后根据相邻帧间预测子块的 MV 推算当前帧间预测子块的运动向量预测值（Motion Vector Prediction，MVP），最后由式(2)计算得到的 MVD 将替代 MV 经过指数哥伦布编码后被写进编码后的码流中。

$$MVD=MV-MVP \tag{2}$$

利用符号相反但绝对值相同的两个数经过指数哥伦布编码后码字字长相同的特性，按照式(3)取 1 个随机比特对 MVD 符号进行随机置乱，即可在保证码率不变的同时影响视频的帧间参考信息。

$$\begin{cases} MVD_{\text{new}}=MVD_{\text{org}} \text{ if } b=0 \\ MVD_{\text{new}}=-MVD_{\text{org}} \text{ if } b=1 \end{cases} \tag{3}$$

其中，$MVD_{\text{org}}^{\text{sign}}$ 和 $MVD_{\text{new}}^{\text{sign}}$ 分别表示加密前后的 MVD，b 为随机比特。

上述两种方法可在不影响码率的同时实现对帧内预测帧（I 帧）和帧间预测帧（P 帧和 B 帧）的有效加密，但不能抵抗最近提出的基于宏块编码比特数的轮廓攻击。因此，将其与本文提出的加密方法相结合可进一步提高视频的安全性。

3　算法设计

3.1　宏块的编码信息结构

图 2 给出了 H.264/AVC 编码标准中宏块编码信息结构，

其包含 5 个部分：mb_type，mb_pred，CBP，QP 和 residual。

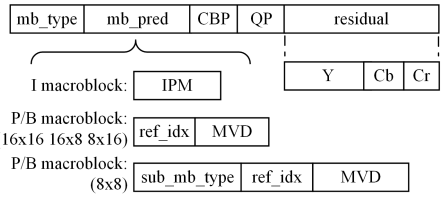


图 2 H.264/AVC 宏块的编码信息

Fig. 2 Coding information of macroblock in H.264/AVC

图 2 中，mb_type 为宏块类型，记录宏块的预测模式和子块划分。mb_pred 包含了宏块内所有与预测相关的信息。具体来说，对于 I 宏块，其包含了各个预测子块的 IPM，对于分割为 16×16，16×8，8×16 的 P 宏块或 B 宏块，其包含了各个预测子块的参考帧索引(ref_idx)和 MVD。而对于分割为 8×8 的 P 宏块或 B 宏块，mb_pred 不仅包含了 ref_idx 和 MVD，还包含了其各个预测子块的类型(sub_mb_type)。CBP 为残差编码方案，用 6 bit 标志位分别表示宏块的亮度和色度是否含有非零的 QDCT 系数。QP 与量化步长相关，QP 值越大，量化步长越长，带来的压缩效率越高，但得到的视频画面质量越差，反之亦然。residual 为残差数据，由宏块所有亮度和色度的 QDCT 系数块组成。

mb_pred 依赖于 mb_type，若两个编码成分不匹配，将会造成解码失败。例如，若 mb_type 表示 Intra_4×4，则 mb_type 应包含 16 个预测子块的 IPM 信息；若 mb_type 表示分割为 8×16 的单向帧间预测类型(P_8×16)，则 mb_pred 应包含两个预测子块的 ref_idx 和 MVD。CBP 必须与 residual 相对应，否则将会导致编码数据丢失或解码失败。特别地，对于 Intra_16×16 类型宏块，其 CBP 包含于 mb_type 中，且 16 个 4×4 的 QDCT 系数矩阵中的 DC 系数需要进一步进行哈达玛(hadamard)变换形成一个额外的系数矩阵；对于 P_Skip 和 B_Skip 类型宏块，它们的像素重建完全依赖于相邻宏块的信息，因此不携带自身数据；对于 B_Direct 类型宏块，它们不携带任何帧间参考的预测信息，由其相邻宏块的预测信息推导而来。

综上，我们将非 Intra_16×16，P_Skip，B_Skip 和 B_Direct 类型宏块的编码信息划分为两个部分：{mb_type，mb_pred，QP}和{CBP，residual}，并将其分别标记为 tpq 和 cr。各个宏块内的 tpq 与 cr 相互独立，这意味着即使这些宏块相互交换其 tpq 或 cr 部分，仍然可以保证视频数据的完整性和编解码格式的兼容性。

3.2 加密算法流程

本文提出的加密算法框图如图 3 所示。首先本文算法需要得到一帧中所有宏块的编码信息，它可由已经编码好的 H.264/AVC 数据经过熵解码得到，也可通过对原始视频像素数据进行除熵编码外所有 H.264/AVC 编码步骤而来。然后，利用帧编码特征自适应产生的 IV 和用户密钥 key，通过 AES-OFB 算法产生加密过程中需要的伪随机序列。经过 IPM 异或、MVD 符号置乱、CBP 和 residual 置换 3 种加密操作后，各宏块进行熵编码输出加密后的 H.264/AVC 码流。

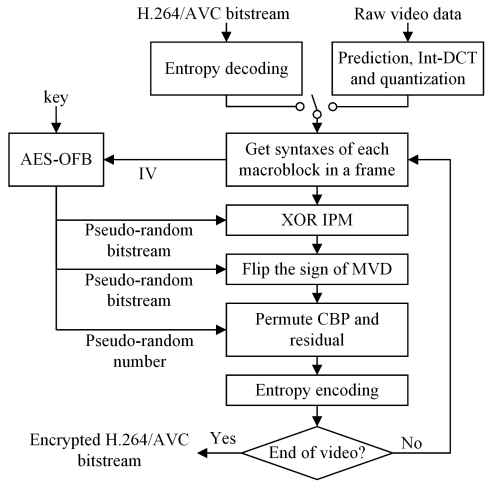


图 3 加密算法框图

Fig. 3 Framework of the proposed encryption algorithm

3.2.1 自适应生成伪随机序列

AES 作为高安全性加密算法[19]，被广泛应用在各种互联网通信与数据存储当中。本文采用 AES 算法的 OFB 模式作为伪随机序列生成器，如式(4)所示：

$$G=AES(iv,key,n)=\{g_i \mid 1\leqslant i\leqslant n\} \tag{4}$$

其中，参数 iv 为 128 bit 的 IV，key 为用户密钥，n 为期望得到的密文分组的个数，返回值 G 表示得到的 n 个 128 bit 长度的密文分组。

对于 AES 加密算法，重复使用 IV 容易遭受重放攻击[20]。为了不使用额外的数据传递 IV，且使得各帧的 IV 不同，本文基于视频帧的特征自适应生成 IV，并保证加密前后该特征不变。由于各宏块的 mb_type 不仅与视频纹理内容相关，也依赖于编码器的参数配置，能较好地反映编码后视频帧的特征，如图 4 所示，针对每一个视频帧，我们按顺序拼接若干宏块的 mb_type 字段经过二值化后的码字，形成 mb_type 二值化序列，取其前 128 bit 为加密当前视频帧时 AES 算法所需的 IV，并通过 AES-OFB 加密器生成 n 个密文分组。

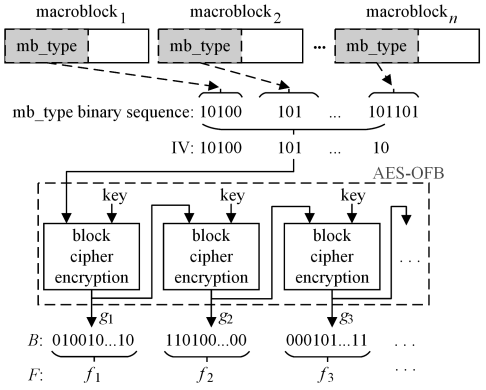


图 4 自适应生成伪随机序列

Fig. 4 Adaptive generate pseudo-random sequence

用于置换加密的伪随机数序列 $F=\{f_i \mid 1\leqslant i\leqslant n\}$ 是各密文分组的十进制表示，其中 $f_i\in[0,2^{128}-1]$，与密文分组一一对应且个数相同，可用式(5)进行描述。

$$f_i=Dec(g_i) \tag{5}$$

用于 IPM 加密和 MVD 符号加密的伪随机比特序列 $B=\{b_i|b_i\in\{0,1\},1\leqslant i\leqslant 128\times n\}$ 由各伪随机数的二进制表示拼接而来,可由式(6)进行计算。

$$b_i=\left(f_{\frac{i-1}{128}+1}\gg(127-(i-1)\bmod 128)\right)\ \&\ 1 \tag{6}$$

其中,≫为右位移运算,& 为与运算,mod 为求模运算。

3.2.2　宏块编码信息置换

对于每一个视频帧,统计其非 Intra_16×16、P_Skip,B_Skip 和 B_Direct 类型宏块的个数,记录为 Q,并将这些宏块的编码信息以 (tpq,cr) 对的形式表示为 $\{(tpq_i,cr_i)|1\leqslant i\leqslant Q\}$。同时,根据式(4)和式(5)自适应生成 Q 个伪随机数 $\{f_i|1\leqslant i\leqslant Q\}$,并将它们一一分配给这些宏块的 cr 部分,表示为 $\{f_i\leftrightarrow cr_i|1\leqslant i\leqslant Q\}$。接着所有宏块的 cr 部分根据伪随机数进行非递减的稳定性值排序(这里的稳定性指若存在多个值相同的元素,那么排序前后它们的相对位置始终保持不变),得到 $\{f_i'\leftrightarrow cr_i'|1\leqslant i\leqslant Q\}$。其中,$\{f_i'|1\leqslant i\leqslant Q\}$ 表示排序后的伪随机数序列,$\{cr_i'|1\leqslant i\leqslant Q\}$ 为跟随伪随机数一起排列后的 cr 序列。最后将置换后的 cr 序列与宏块原 tpq 序列重新配对,形成的新宏块编码信息表示为 $\{(tpq_i,cr_i')|1\leqslant i\leqslant Q\}$。

如图 5 所示,5 个宏块编码信息分别表示为 $\{(tpq_1,cr_1),(tpq_2,cr_2),(tpq_3,cr_3),(tpq_4,cr_4),(tpq_5,cr_5)\}$,对其生成伪随机数并分配给每个宏块的 cr 部分,有 $\{4\leftrightarrow cr_1,2\leftrightarrow cr_2,3\leftrightarrow cr_3,1\leftrightarrow cr_4,2\leftrightarrow cr_5\}$,基于伪随机数排序后得到的新序列为 $\{1\leftrightarrow cr_4,2\leftrightarrow cr_2,2\leftrightarrow cr_5,3\leftrightarrow cr_3,4\leftrightarrow cr_1\}$,因此,置换加密后新配对的宏块编码信息组成为 $\{(tpq_1,cr_4),(tpq_2,cr_2),(tpq_3,cr_5),(tpq_4,cr_3),(tpq_5,cr_1)\}$。

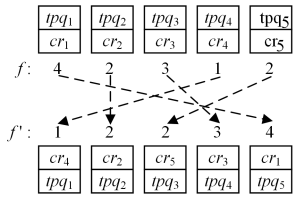


图 5　置换残差编码方案和残差数据

Fig. 5　Permute CBP and residual

4　实验结果与分析

本文的加密方法基于开源编码器 JM19.0[1)] 实现,共选取了来自 3 个数据集(Xiph,ICDAR2013[21] 和 MCL-JCV[22])的 9 个测试视频序列,具体如表 1 所列,每个视频序列只选取前 30 帧数据进行测试。这些视频包含了不同的场景、主体、纹理、色彩(灰度)、分辨率等,能够从多个角度较好地测试加密算法的优劣。视频序列的 GOP 结构为"IBPBP",这意味着每个视频序列包含 6 个 GOP。如果没有特殊说明,实验使用的默认 QP 值为 28。

表 1　测试视频描述

Table 1　Description of tested video

Dataset	Video	Resolution
Xiph	soccer	704×576
	pedestrian	1 920×1 080
	tractor	1 920×1 080
ICDAR2013	Video11	1 280×960
	Video17	1 280×960
	Video20	1 280×960
MCL-JCV	videoSRC02	1 280×720
	videoSRC11	1 280×720
	videoSRC19	1 280×720

4.1　视觉安全性

结构相似性(Structural Similarity,SSIM)[23] 和视频多方法评估融合(Video Multimethod Assessment Fusion,VMAF)[24] 这两个指标被广泛应用于视频加密算法的视觉安全性评估,其中 $SSIM\in[0,1]$,$VMAF\in[0,100]$,值越小,加密视觉效果越好,反之亦然。表 2 列出了各视频序列分别在 QP 值为 20,28 和 36 这 3 种情况下,Ding 等[9]、Xu 等[14]、Liu 等[16] 提出的加密算法和本文(Ours)加密算法的测试值。相比未经加密的视频(Orig),4 种加密方法在两个指标上显著降低,且在绝大多数情况下,本文提出的加密算法优于 Ding 等[9],Xu 等[14],Liu 等[16] 提出的加密算法。

表 2　视觉安全性指标

Fig. 2　Visual safety indicators

video	QP	*VMAF*					*SSIM*				
		Orig	Ding 等[9]	Xu 等[14]	Liu 等[16]	Ours	Orig	Ding 等[9]	Xu 等[14]	Liu 等[16]	Ours
soccer	20	99.846	10.515	7.353	10.276	**6.482**	0.994	0.143	0.146	0.146	**0.113**
	28	95.251	9.848	7.285	9.854	**6.495**	0.973	0.169	0.205	0.154	**0.123**
	36	77.204	9.522	6.803	9.712	**5.333**	0.895	0.204	0.242	0.219	**0.203**
pedestrain	20	99.807	15.362	12.494	14.941	**10.525**	0.998	0.244	0.252	0.260	**0.227**
	28	96.987	13.658	12.229	13.641	**8.245**	0.993	0.276	0.261	**0.250**	0.282
	36	77.160	14.087	13.089	15.002	**7.126**	0.976	0.289	0.302	**0.261**	0.355
tractor	20	99.848	14.916	13.122	14.845	**11.438**	0.999	0.119	0.138	0.122	**0.108**
	28	99.512	14.326	12.535	14.143	**11.460**	0.994	0.127	0.157	0.131	**0.117**
	36	81.924	14.641	12.821	14.586	**10.808**	0.974	0.136	0.154	**0.133**	0.139
video11	20	99.773	6.224	2.907	6.415	**2.167**	0.998	**0.319**	0.353	0.340	0.321
	28	98.222	7.504	2.997	6.481	**2.586**	0.992	0.390	0.367	0.354	**0.347**
	36	87.442	6.079	**3.604**	6.349	3.607	0.971	**0.306**	0.344	0.360	0.374
video17	20	99.835	13.687	11.444	13.406	**10.110**	0.999	0.249	0.253	0.244	**0.184**
	28	99.647	13.846	11.570	13.173	**9.687**	0.992	0.255	0.288	0.245	**0.214**
	36	86.234	12.973	11.628	13.996	**8.267**	0.965	**0.263**	0.283	0.268	0.269

1) http://ipome.hhi.de/suehring/tm1/

（续表）

video	QP	VMAF					SSIM				
		Orig	Ding 等[9]	Xu 等[14]	Liu 等[16]	Ours	Orig	Ding 等[9]	Xu 等[14]	Liu 等[16]	Ours
video20	20	99.865	5.872	1.707	5.524	**0.828**	0.998	0.312	0.256	0.270	**0.199**
	28	99.626	6.086	2.094	5.569	**1.058**	0.992	0.266	0.266	0.267	**0.232**
	36	81.597	5.591	2.237	5.763	**0.591**	0.965	0.267	0.299	0.278	**0.260**
videoSRC02	20	99.837	10.957	4.749	8.962	**4.714**	0.998	0.271	0.284	0.282	**0.248**
	28	97.095	13.005	5.853	8.795	**3.985**	0.992	0.271	0.290	0.280	**0.265**
	36	77.787	11.340	6.685	12.056	**3.102**	0.971	**0.269**	0.294	0.290	0.296
videoSRC11	20	99.839	6.164	3.863	6.164	**2.529**	0.997	0.275	0.249	0.316	**0.217**
	28	98.184	7.178	4.048	6.986	**1.440**	0.990	0.281	0.294	0.279	**0.243**
	36	78.972	7.420	4.359	7.345	**1.122**	0.965	0.296	0.328	0.281	**0.280**
videoSRC19	20	99.842	12.372	**7.667**	12.114	7.697	0.997	0.199	0.222	0.183	**0.176**
	28	98.481	11.445	8.178	10.814	**6.114**	0.983	0.219	0.244	0.221	**0.194**
	36	77.644	12.041	8.724	9.867	**4.684**	0.936	0.244	0.276	**0.243**	0.255

图 6 给出了使用本文加密算法加密后的视频画面。

(a) tractor original-1-I (b) tractor original-15-P (c) tractor original-24-B

(d) tractor encrypted-1-I (e) tractor encrypted-15-P (f) tractor encrypted-24-B

注：1/15/24 分别表示帧编号，I/P/B 分别表示帧类型

图 6 加密视频的图像

Fig. 6 Images of encrypted videos

相比原始视频，无论是从色彩（灰度）还是内容上，我们都无法从加密后的视频画面中感知到任何有效信息。值得一提的是，I 帧的视觉加密效果会通过帧间预测传播到同一 GOP 中的其他 P 帧和 B 帧，因此相对于 I 帧，P 帧和 B 帧的视觉加密效果通常更好。此外，MVD 符号的改变导致 MV 重建错误，进一步造成错误的宏块帧间参考，因此在 P 帧和 B 帧中会出现块失真效应（见图 6 (e)），且对于运动越剧烈的视频序列，块失真效应越明显。

4.2 轮廓攻击

在视频编解码中，宏块纹理越复杂，编码该宏块所需的比特数就越多。Minemura 等[17]根据该特性，统计一帧中各宏块的比特开销，通过式(7)构造加密视频的轮廓攻击图。

$$\phi(i,j)=round\left(255\times\frac{c(i,j)}{\max\{c(i,j)\}}\right) \tag{7}$$

其中，$c(i,j)$为编码帧中第(i,j)个宏块的比特开销，$\phi(i,j)$为轮廓攻击图中对应的第(i,j)个像素点的值，$round(x)$为向下取整函数，$\max\{\cdot\}$表示返回最大值。由于编码帧的一个宏块对应轮廓攻击图的一个像素点，因此轮廓攻击图的长宽分辨率各是原视频图像的 1/16。图 7 给出了各加密算法的轮廓攻击图，其结果表明 Ding 等[9]、Xu 等[14]、Liu 等[16]的加密算法易遭受轮廓攻击，而本文方法能够有效扰乱所生成的轮廓。

(a) videoSRC02 Orig (b) videoSRC02 Ding 等[9] (c) videoSRC02 Xu 等[14] (d) videoSRC02 Liu 等[16] (e) videoSRC02 Ours

(f) videoSRC11 Orig (g) videoSRC11 Ding 等[9] (h) videoSRC11 Xu 等[14] (i) videoSRC11 Liu 等[16] (j) videoSRC11 Ours

图 7 各加密方法的轮廓攻击图

Fig. 7 Outline images of videos encrypted by each method

边缘相似性分数（Edge Similarity Score，ESS）被用来评价两幅边缘图像的相似度，一幅是式(7)计算得到的轮廓攻击图 ϕ，另一幅为视频原图经过边缘检测后得到的标准边缘图，标记为 S。一方面，由于 S 的长宽分辨率是 ϕ 的 16 倍，因此将 S 分割为无重叠的 16×16 子块，并通过式(8)进行下采样二值化。

$$S'(i,j)=\begin{cases}1, & \text{if } \sum_{i=u}^{16}\sum_{j=v}^{16}S_{u,v}(i,j)>0\\ 0, & \text{otherwise}\end{cases} \tag{8}$$

其中，$S'(i,j)$为标准边缘图下采样二值化后的第(i,j)个像素点的值，$S_{u,v}(i,j)$是标准边缘图中第(i,j)个子块的第(u,v)个像素点的值。另一方面，轮廓攻击图同样需要进行二值化，得到的图像标记为 ϕ'。最后，ESS 指标可由式(9)计算：

$$ESS=\left|\frac{n_1}{2Z}+\frac{n_2}{2O}-\frac{n_3}{2Z}-\frac{n_4}{2O}\right| \tag{9}$$

其中，Z 和 O 分别表示 ϕ' 中值为 0 和 1 的像素点个数，$n_i(i=1,2,3,4)$分别由下式定义：

$$
\begin{aligned}
n_1 &= |\{(i,j)|S'(i,j)=0 \text{ and } \phi'(i,j)=0\}| \\
n_2 &= |\{(i,j)|S'(i,j)=1 \text{ and } \phi'(i,j)=1\}| \\
n_3 &= |\{(i,j)|S'(i,j)=0 \text{ and } \phi'(i,j)=1\}| \\
n_4 &= |\{(i,j)|S'(i,j)=1 \text{ and } \phi'(i,j)=0\}|
\end{aligned} \tag{10}
$$

ESS 的取值范围为[0,1]，分数越高，轮廓攻击的效果越好，反之亦然。实验中，我们使用 Canny 边缘检测器[25]获取标准边缘图像，并通过 OTSU 算法[26]对轮廓攻击图进行二值化，计算得到的 *ESS* 指标展示在表 3 中。可以得出，本文的方法在抵抗轮廓攻击方面具有较为明显的优势。

表 3　边缘相似性分数、加密空间和像素变化率

Table 3　*ESS*, encryption space and *NPCR*

Video	*ESS*					Encryption Space			*NPCR*/%
	Orig	Ding 等[9]	Xu 等[14]	Liu 等[16]	Ours	I frame	P frame	B frame	
soccer	0.105	0.104	0.105	0.104	**0.077**	1.185×10^{4068}	3.091×10^{3164}	5.386×10^{903}	99.403
pedestrian	0.386	0.385	0.386	0.385	**0.318**	2.637×10^{14971}	8.993×10^{12767}	3.369×10^{7939}	99.524
tractor	0.234	0.234	0.234	0.234	**0.161**	3.711×10^{24969}	5.172×10^{20929}	1.152×10^{11530}	99.513
video11	0.334	0.333	0.334	0.333	**0.214**	2.162×10^{8812}	3.025×10^{9239}	7.090×10^{7015}	98.897
video17	0.281	0.281	0.281	0.281	**0.150**	3.266×10^{12271}	1.073×10^{11720}	1.129×10^{8840}	99.508
video20	0.552	0.552	0.552	0.552	**0.296**	2.437×10^{10618}	5.558×10^{10293}	1.085×10^{7272}	98.989
videoSRC02	0.344	0.344	0.344	0.345	**0.291**	6.636×10^{5617}	1.048×10^{4621}	4.476×10^{2282}	96.397
videoSRC11	0.583	0.583	0.583	0.583	**0.425**	4.737×10^{5397}	1.526×10^{4876}	3.672×10^{2858}	96.469
videoSRC19	0.176	0.177	0.176	0.176	**0.130**	1.472×10^{9438}	1.036×10^{7053}	3.339×10^{3777}	96.679

4.3　加密空间

加密空间的大小是评价加密算法能否抵抗暴力攻击的参考标准。对于宏块间的 CBP 和 residual 置换，由于进行置换的宏块的 residual 含有 16 个 4×4 的 QDCT 系数块，各系数块又包含 16 个 QDCT 系数，且系数取值多样，因此可将每个置换的对象视为唯一，则置换为全排列方式，加密空间可由式(11)计算：

$$space = Q! \tag{11}$$

其中，Q 为一帧中非 Intra_16×16，P_Skip，B_Skip 和 B_Direct 类型宏块的个数。表 3 记录了实际编码时各个视频序列不同帧类型的置换加密空间的中位数。可以发现，相比 I 帧，P 帧的置换加密空间较小，B 帧最小。由于 P 帧和 B 帧中的宏块可以采用帧间预测进行编码，相比 I 帧中宏块的帧内预测模式，帧间预测往往准确性更高，带来更少的 QDCT 系数，进而更多宏块被编码成 P_Skip，B_Skip 和 B_Direct 类型。因此，P 帧和 B 帧中的 Q 值相比 I 帧往往小很多。然而，由于 P 帧和 B 帧进行帧间参考的特点，它们的正确解密需要基于其先序的所有参考帧的正确解密，因此安全性更高。

4.4　密钥敏感性

密钥敏感性反映密钥在只有很小变化的情况下对加密效果的影响，是证明加密算法能够有效抵抗差分攻击的判定标准之一。像素变化率(Number of Changing Pixel Rate，NPCR)在图像和视频领域常用来评价加密视觉效果对密钥的敏感性[27]。NPCR 由式(12)和式(13)计算，其在 8 bit 位深视频下的理论最优值为 99.609%，算法越接近理论最优值，其密钥敏感性越好。

$$D(i,j)=\begin{cases}0, & \text{if } P_1(i,j)=P_2(i,j)\\ 1, & \text{if } P_1(i,j)\neq P_2(i,j)\end{cases} \tag{12}$$

$$NPCR=\frac{1}{h\times w}\times D(i,j)\times 100\% \tag{13}$$

其中，h 和 w 为视频帧的高和宽，$P_1(i,j)$ 和 $P_2(i,j)$ 分别表示使用两个只有 1bit 差异的密钥加密得到的两个视频帧的第 (i,j) 个像素点的值。表 3 列出了各视频序列以帧为单位的 *NPCR* 平均值，可以看出，大部分数字逼近于理论最优值，表明了本文的加密算法拥有良好的密钥敏感性。

4.5　码率变化

加密是一个熵增的过程，这违背了视频编码中去冗余的特点。若加密操作导致视频压缩效率变低、码率急剧上升，则会带来大量的带宽负担和存储成本，因此将码率变化控制在一定范围内具有极其重要的实际应用价值。码率变化的计算方式如下：

$$inc=\frac{BR'-BR}{BR} \tag{14}$$

其中，BR 和 BR' 分别为加密前后视频的码率。图 8 给出了部分视频序列在不同 *QP* 值下加密后的码率增量。随着 *QP* 值的增加，码率增量大致呈现下降的总趋势。这表明本文算法能够较好地控制码率的增长(大部分维持在 0.8%以下)。

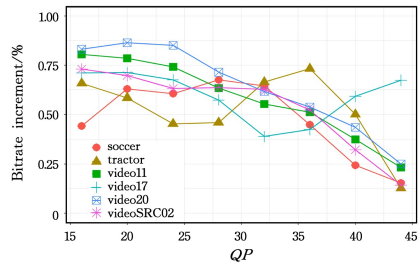


图 8　加密视频在不同 *QP* 值下的码率增量

Fig. 8　Bitrate increment of each encrypted video with different *QP* values

造成码率增长的主要原因有两点：1)由于宏块间进行了 residual 置换，破坏了 QDCT 系数块之间的相关性，造成熵编码时的额外开销；2)在上下文自适应变长编码(Context-based Adaptive Variable Length Coding，CAVLC)中，I 宏块的 CBP 码字映射表与 P 宏块和 B 宏块的不同，当 P 宏块或 B 宏块与 I 宏块相互交换 CBP 值时，将导致 CBP 码字长度发生变化。

结束语　视频隐私保护对高速发展的云技术至关重要。本文提出了一种基于宏块编码信息置换的 H.264/AVC 视频加密方法，其与现有的选择加密方法相比拥有更好的视觉安全性，并能抵抗轮廓攻击。此外，该方法还拥有加密空间大、密钥敏感性好、对视频码率影响小等特点。未来我们将会进一步针对新一代视频编解码标准(如 H.265/HEVC)提出一套可行的隐私保护方案。

参 考 文 献

[1] TABASH F K, IZHARUDDIN M. Encryption techniques for H. 264/AVC videos: A literature review[J]. Journal of Information Security and Applications, 2019, 45(APR.): 20-34.

[2] BOHO A, WALLENDAEL G V, DOOMS A, et al. End-to-end security for video distribution[J]. IEEE Signal Processing Magazine, 2013, 30(2): 97-107.

[3] AHN J, SHIM H J, JEON B, et al. Digital VideoScrambling Method Using Intra Prediction Mode[C]// Conference on Advances in Multimedia Information Processing-PCM. 2004: 386-393.

[4] LI Y, LIANG L, SU Z, et al. A New Video Encryption Algorithm for H. 264[C]// IEEE International Conference on Information, 2005: 1121-1124.

[5] KHLIF N, DAMAK T, KAMMOUN F, et al. Motion vectors signs encryption for H. 264/AVC[C]// IEEE International Conference on Advanced Technologies for Signal & Image Processing, 2014: 1-6.

[6] SU P C, HSU C W, WU C Y. A practical design of content protection for H. 264/AVC compressed videos by selective encryption and fingerprinting[J]. Multimedia Tools & Applications, 2011, 52(2/3): 529-549.

[7] SHEN H, ZHUO L, ZHAO Y. An efficient motion reference structure based selective encryption algorithm for H. 264 videos [J]. IET Information Security, 2014, 8(3): 199-206.

[8] WANG Y, O'NEILL M, KURUGOLLU F. Partial encryption by randomized zig-zag scanning for video encoding[C]// IEEE International Symposium on Circuits & Systems. 2013: 229-232.

[9] DING X, DENG Y, YANG G, et al. Design of new scan orders for perceptual encryption of H. 264/AVC videos[J]. IET Information Security, 2017, 11(2): 55-65.

[10] DI X Q, WANG Y Z, LI J Q, et al. Video encryption method based on hyperchaos of quantum cellular neural network[J]. Journal of Jilin University (Engineering and Technology Edition), 2018, 48(3): 919-928.

[11] BAI S, GUO Y, ZHAO B, et al. H. 264 video perceptual encryption algorithm with controllable visual quality based on CABAC [J]. Journal of Electronics and Information, 2016, 38(10): 2582-2589.

[12] WANG Y S, O'NEILL. A Tunable Encryption Scheme and Analysis of Fast Selective Encryption for CAVLC and CABAC in H. 264/AVC[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2013, 23(9): 1476-1490.

[13] PENG F, GONG X Q, LONG M, et al. A selective encryption scheme for protecting H. 264/AVC video in multimedia social network[J]. Multimedia Tools and Applications, 2017, 76(3): 3235-3253.

[14] XU D, WANG R, SHI Y Q. Data Hiding in Encrypted H. 264/AVC Video Streams by Codeword Substitution[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(4): 596-606.

[15] KHLIF N, MASMOUDI A, KAMMOUN F, et al. Secure chaotic dual encryption scheme for H. 264/AVC video conferencing protection[J]. IET Image Processing, 2018, 12(1): 42-52.

[16] LIU S, RHO S, JIFARA W, et al. A hybrid framework of data hiding and encryption in H. 264/SVC[J]. Discrete Applied Mathematics, 2018, 241: 48-57.

[17] MINEMURA K, WONG K. A novel sketch attack for H. 264/AVC format-compliant encrypted video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(11): 2309-2321.

[18] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H. 264/AVC video coding standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2003, 13(7): 560-576.

[19] LIU B, BAAS B M. Parallel AES Encryption Engines for Many-Core Processor Arrays[J]. IEEE Transactions on Computers, 2013, 62(3): 536-547.

[20] SKARLATOS D, YAN M, GOPIREDDY B, et al. MicroScope: Enabling Microarchitectural Replay Attacks[C]// ACM/IEEE 46th Annual International Symposium on Computer Architecture (ISCA). 2019: 318-331.

[21] KARATZAS D, SHAFAIT F, UCHIDA S, et al. ICDAR 2013 robust reading competition[C]// 12th International Conference on Document Analysis and Recognition (ICDAR). 2013: 1484-1493.

[22] WANG H, GAN W, HU S, et al. MCL-JCV: A JND-based H. 264/AVC video quality assessment dataset[C]// IEEE International Conference on Image Processing (ICIP). 2016: 1509-1513.

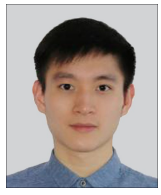
[23] OU T S, HUANG Y H, CHEN H H. SSIM-Based Perceptual Rate Control for Video Coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(5): 682-691.

[24] LI Z, AARON A, KATSAVOUNIDISI, et al. Toward A Practical Perceptual Video Quality Metric[EB/OL]. (2016-06-06). http://techblog.netflix.com/2016/06/toward-practical-perceptual-video.html.

[25] CANNY J. A Computational Approach To Edge Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6): 679-698.

[26] OTSU N. A threshold selection method from gray level histogram[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.

[27] WU Y, NOONAN J S, AGAIAN S. NPCR and UACI Randomness Tests for Image Encryption[J]. Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications, 2011(April): 31-38.



LIANG Jian, born in 1996, postgraduate. His main research interests include cyber security and multimedia signal processing.

HE Jun-hui, born in 1976, Ph. D, associate professor. His main research interests include cyber security and multimedia signal processing.

(责任编辑:柯颖)