融合视频编码的低复杂度纹理自适应视频加密算法

刘慧超[1,2]，王志君[1*]，梁利平[1]

(1. 中国科学院微电子研究所 北京 朝阳区 100029；2. 中国科学院大学电子电气与通信工程学院 北京 石景山区 100049)

【摘要】针对现有视频加密算法复杂度高的问题，提出一种低复杂度的纹理自适应视频加密算法。该算法直接利用视频编码 DCT 系数和运动矢量建立阈值自适应的纹理估计模型，以此检测视频纹理区域并进行加密。考虑到视频编码效率和格式兼容性，选择纹理区域的 DCT 系数符号和 MVD 符号进行加密。该算法以应用广泛的 H.264/AVC 视频编解码器为验证平台，实验结果表明，该文提出的加密算法在确保视频内容安全性和压缩编码效率的同时，加密算法复杂度降低，待加密数据量减少。与现有纹理检测模型相比，该文纹理检测算法复杂度低，能够适用于实时性视频加密应用场景。

关 键 词 数字加密； 选择性加密； 纹理检测； 视频压缩； 视频加密

中图分类号 TN919.3 **文献标志码** A doi:10.12178/1001-0548.2019291

Low Complexity Texture-adaptive Video Encryption Algorithm Fused with Video Coding

LIU Hui-chao[1,2], WANG Zhi-jun[1*], and LIANG Li-ping[1]

(1. Institute of Microelectronics, Chinese Academy of Sciences Chaoyang Beijing 100029;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences Shijingshan Beijing 100049)

Abstract Aiming at the problem of high complexity of existing video encryption algorithms, a low complexity texture-adaptive video encryption algorithm is proposed. Firstly, the quantized residual DCT coefficients and motion vectors obtained in the video encoding process are used to build the threshold-adaptive texture estimation model, which is used to detect texture region for video encryption. Then, considering the video encoding efficiency and format compatibility, the signs of quantized non-zero DCT coefficient and non-zero MVD are selected for encryption in the texture region, respectively. The algorithm takes the H.264/AVC video codec as a verification example, the experimental results show that, the proposed encryption algorithm ensures the security of video content and the efficiency of compression encoding, while reducing the complexity of the encryption algorithm and the amount of data to be encrypted greatly. Meanwhile, compared with existing texture detection methods, the complexity of the proposed texture detection algorithm is extremely low, and it can be applied to real-time video encryption application scenarios.

Key words digital encryption; selective encryption; texture detection; video compression; video encryption

随着互联网技术的飞速发展，基于视频信息的多媒体应用越来越普遍化、多样化，对于视频数据内容的加密保护变得愈发重要[1-2]，数字视频加密技术应运而生。传统方案利用 AES、DES 等经典加密算法对原始视频数据进行加密。然而，庞大的视频数据导致加密过程实时性较差；与此同时，加密后的视频数据杂乱无章，直接影响后续视频压缩编码的效率，因而此类方法并不适用于视频加密。相应地，针对视频压缩码流的加密算法导致码流格式不兼容，通用解码器无法正常解码。因此，与视频压缩编码过程相融合的选择性加密算法 (selective encryption, SE) 成为新的研究热点[3-4]。

SE 算法通过对视频压缩编码过程中的关键句法进行加密实现内容保护。视频编码过程中，I 帧作为每个编码序列的起始帧包含了视频序列的大部分信息，文献 [5-7] 提出对 I 帧编码块的帧内预测

收稿日期：2019－12－19；修回日期：2020－03－01

作者简介：刘慧超 (1991－)，男，博士生，主要从事视频编解码、视频加密、视频水印和媒体处理器等方面的研究.

通信作者：王志君，E-mail：wangzhijun@ime.ac.cn

模式 (intra prediction mode, IPM) 进行加密，然而该方法易导致加密后的 IPM 不再满足编码规范，解码端无法正常解码；相比于 I 帧，P/B 帧在每个编码序列中出现最频繁，文献 [8−10] 提出加密 P/B 帧的运动矢量来保护视频内容，然而视频编码算法对运动矢量 (motion vector, MV) 采用预测编码，因此，此类方法将直接改变句法元素 MVD 的大小，从而影响视频压缩效率。为增强加密效果，对 IPM、MVD 符号和 DCT 系数等多种句法元素的联合加密方案相继被提出[11-17]。文献 [13-14] 仅对 MVD 符号位、非零 DCT 系数符号位以及系数幅值的二进制编码后缀进行加密，编码效率恒定，但加密数据量明显增加。为此，文献 [18] 提出采用混沌系统随机选择待加密句法元素，但无法保证视频关键区域绝对保密。于是，文献 [19-21] 融合感兴趣区域 (regions of interest, ROI) 技术对视频关键区域进行选择性加密，既降低了加密数据量，同时视频关键区域信息也得到了保护，然而却引入了如高斯混合模型 (gaussian mixture model, GMM) 等计算复杂的 ROI 检测算法，难以做到实时性。

针对以上问题，本文提出一种基于视频纹理特性的低复杂度选择性加密算法。算法以应用广泛的 H.264/AVC 视频编码标准为验证实例，首先利用量化的残差 DCT 系数和运动矢量建立低复杂度的视频纹理强度估计模型；然后根据纹理强度对 I 帧编码宏块的非零 DCT 系数符号位和 P/B 帧编码宏块的 MVD 符号位进行自适应选择加密，实现格式兼容、压缩效率恒定、复杂度低、加密效果好的视频加密。此外，提出一种与视频纹理特征相关的密钥生成算法，密钥安全性进一步提高。

1　融合加密技术的视频混合编码架构

自 H.261 视频编码标准起，新标准在引入先进技术的同时，一直沿用基于预测和变换的混合编码架构，如图 1 所示，混合编码器中主要有两条路径：前向路径和重建路径。其中，前向路径包括预测编码、DCT 变换与量化和熵编码。预测编码利用相邻块的空间或时间相关性，根据已编码块对当前待编码块进行线性预测，然后对预测值和真值的差进行编码，减小待编码数据量；变换编码则将统计上彼此密切相关的空域像素通过正交变换，转化为统计上相对独立的一系列变换系数，减小有效数据量。熵编码过程通常为变长编码，通过对出现概率大的字符分配短码字、概率小的字符分配长码字，进一步提高混合编码的压缩效率。重建路径主要包括反变换与反量化、环路滤波。反变换与反量化得到重构残差，与预测值叠加得到重建像素，之后经环路滤波得到重建帧，作为后续编码过程的参考帧。如图 1 所示为混合编码核心结构框图。

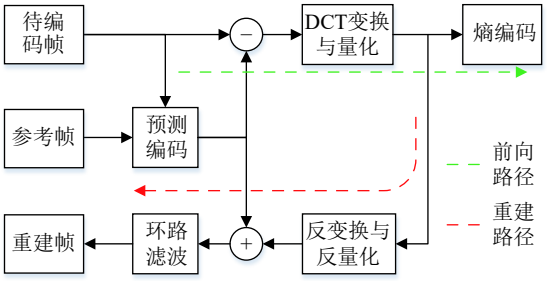


图 1　混合编码器结构框图

结构上，加密模块在编码器中的位置直接影响编码端算法复杂度。若将加密过程设置在重建环路内，则在编码器中需要配置对应的解密过程，增加计算复杂度。因此，本文将加密模块设置在重建环路外，如图 2 所示，加密模块位于 DCT 变换与量化之后、熵编码之前的前向编码路径中。

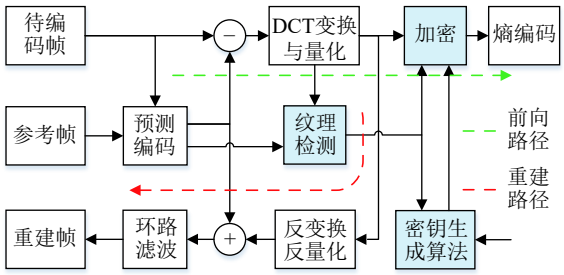


图 2　融合加密的视频混合编码器结构框图

图 2 中，纹理检测模块直接利用 DCT 变换和量化模块输出的量化 DCT 系数与预测编码模块输出的 MV 进行纹理估计，根据设定的纹理强度阈值，将每个视频帧划分为纹理和非纹理区域，视频加密模块根据划分结果对纹理区域执行选择性加密过程。

2　基于纹理的视频加密算法

视频加密旨在扰乱视频内容，使得人眼难以分辨出视频画面所传达出的有效信息，从而实现视频内容保护。文献 [22] 通过对视频编码句法元素进行下采样，实现全局上的部分加密。然而，根据人眼视觉特性，人眼系统对视频纹理区域更加感兴趣。于是，视频加密过程只需要处理纹理区域即可，无需对全局视频信息进行全加密。基于此，本文分别提出针对 I 帧和 P 帧 (B 帧与之类似，暂不讨论) 的纹理检测模型，并选择纹理区域进行加

密，有效降低待加密数据量，同时保证视频信息的安全性。

2.1 基于 DCT 的 I 帧纹理检测模型

对于编码 I 帧，即采用帧内预测方式进行编码的视频帧，主要利用多方向线性预测对待编码帧进行估计，预测值与真实值的差，即预测残差，经 DCT 变换和量化后得到一系列相互独立的量化 DCT 系数。若以 $\boldsymbol{X}_o$ 表示原始帧像素集合，$\boldsymbol{X}_p$ 表示最佳预测模式下的预测帧像素集合，$\boldsymbol{X}_r$ 表示残差集合，则有：

$$\boldsymbol{X}_r = \boldsymbol{X}_o - \boldsymbol{X}_p \tag{1}$$

由于 $\boldsymbol{X}_p$ 仅仅是利用 $\boldsymbol{X}_o$ 的相邻边界像素进行线性预测的结果，除平坦区域外，其余区域的线性预测通常情况下均会产生一定的预测残差，且 $\boldsymbol{X}_o$ 纹理越复杂，线性预测产生的残差越显著，对残差进行 DCT 变换和量化得到的系数矩阵包含的非零系数就越多。对于一个 $n\times n$ 的残差块，定义其 DCT 变换和量化后的系数矩阵如下：

$$\boldsymbol{Y}_R = \mathrm{QDCT}(\boldsymbol{X}_R)=\begin{bmatrix} y_{0,0} & y_{0,1} & \cdots & y_{0,n-1} \\ y_{1,0} & y_{1,1} & \cdots & y_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n-1,0} & y_{n-1,1} & \cdots & y_{n-1,n-1} \end{bmatrix} \tag{2}$$

对于系数矩阵 $\boldsymbol{Y}_R$，其包含的非零系数越多，$\boldsymbol{X}_R$ 的时域特性越复杂，纹理越丰富。与此同时，$\boldsymbol{Y}_R$ 中高频位置非零系数越多，$\boldsymbol{X}_R$ 纹理细节越丰富。根据 DCT 变换的性质，对于 $\boldsymbol{Y}_R$，左上角位置系数 $y_{0,0}$ 为直流分量，右下角位置系数 $y_{n-1,\,n-1}$ 为最高频分量，同一行（或列）的不同系数，从左至右（或从上至下）频率逐渐升高。因此，为定量分析纹理特性，本文为 $y_{i,j}$ 定义如下权值系数：

$$w_{i,j} = (i+j) \quad i,j \in [0, n-1] \tag{3}$$

式中，系数 y_{ij} 的权值直接由其所在位置决定，越靠近左上角权值越低，相反，越靠近右下角权值越高，符合 DCT 变换频率分布特点。本文提出一种评估 $\boldsymbol{X}_o$ 纹理强度 (texture strength, TS) 的量化方法，如式 (4) 所示：

$$\mathrm{TS}_I = \sum_i \sum_j (nz_{i,j} * w_{i,j}) \quad nz_{i,j} = \begin{cases} 0 & y_{i,j} = 0 \\ 1 & y_{i,j} \neq 0 \end{cases} \tag{4}$$

设定一个纹理强度阈值 T_I，当 $\mathrm{TS}_I \geqslant T_I$ 时，认为该区域为纹理区域，否则为非纹理区域，如式 (5) 所示：

$$\begin{cases} \mathrm{TS}_I \geqslant T_I & \text{纹理区域} \\ \text{otherwise} & \text{非纹理区域} \end{cases} \tag{5}$$

然而，对于一个视频序列，其内容通常是随时间不断变化的，故而阈值 T_I 也应该随视频内容变化而变化。本文提出一种基于帧平均的 I 帧阈值可伸缩自适应更新算法 (threshold scalable self-adaptive update method for I-frame, TSSUM-I)，其基本思想是根据前一帧的平均纹理强度和伸缩因子更新下一帧的纹理强度阈值，从而动态调整加密强度，相比文献 [23] 提出的基于当前帧被引用频次的动态调节算法和文献 [24] 提出的基于 SSIM 的动态调节算法，本文算法计算复杂度更低。具体实现过程如下：

1) 为第 1 个待编码 I 帧设置初始阈值：

$$T_I(1) = T_{I0} \tag{6}$$

2) 利用式 (4) 计算当前编码 I 帧 (定义为第 n 个 I 帧) 中 (x,y) 位置待编码块的纹理强度 $\mathrm{TS}_I(n,x,y)$，并统计当前编码 I 帧中所有编码单元 (coding units, CU) 的纹理强度和：

$$T_{\mathrm{sum}-I} = \sum_i \sum_j (\mathrm{TS}_I(n,x,y)) \tag{7}$$

3) 根据当前编码 I 帧的平均纹理强度与伸缩因子更新下一个待编码 I 帧 (定义为第 n+1 个 I 帧) 的纹理强度阈值：

$$T_I(n+1) = \lambda_I \mathrm{mean}(T_{\mathrm{sum}-I}) \tag{8}$$

式中，λ_I 为伸缩因子，其值大于零。

4) 若编码尚未完成，则跳转到过程 2)，继续执行相同操作；否则，结束流程。

至此，得到一个基于编码残差量化 DCT 系数的阈值自适应纹理检测模型 (threshold self-adaptive texture detection model based on quantized DCT Coefficient, TSTDM-QDCTC)。为了验证 TSTDM-QDCTC 的有效性，本文以经典的 Canny 边缘检测算法为参考模型，对两种算法的纹理检测结果进行了对比，如图 3 所示。

从图 3 中可以看出，在对纹理检测精度要求不高的场景下，如选择性视频加密，本文提出的 TSTDM-QDCTC 纹理检测模型能够有效的检测出视频纹理区域；同时，根据式 (4) 和式 (5)，对于一个 4×4 的编码块，本文纹理检测算法仅仅需要 15 次加法和 1 次判断，而 Canny 边缘检测计算复杂，折合加法运算约为 3 300 次，计算量远远高于本文算法。

2.2 基于运动矢量的 P 帧纹理检测模型

对于编码 P 帧，即采用帧间预测进行编码的视

频帧，主要利用运动估计和运动补偿技术进行高效率压缩编码，首先通过运动估计得到当前预测单元 (prediction blocks, PU) 相对于其最佳参考块的运动矢量，之后利用插值运算进行运动补偿，得到当前 PU 的最佳预测值，如图 4 所示。

a. 原图

b. Canny 算子检测结果

c. TSTDM-QDCTC 检测结果

图 3　纹理检测结果对比

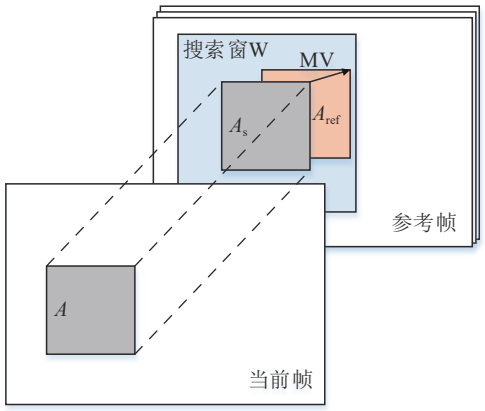


图 4　运动估计示意图

图 4 中，A 为当前待编码 PU，A_s 为参考帧中与 A 位置相对应的已编码 PU，A_{ref} 为在给定搜索窗 W 范围内与 A 匹配最佳的参考区域，A_{ref} 相对于 A_s 的位置坐标即为 A 经运动估计后得到的运动矢量 MV，由水平和垂直两个分量组成，用 mv_x 和 mv_y 表示。mv_x 和 mv_y 直接指出了当前编码 PU 相对于其参考帧的变化程度，于是可以用作待编码 PU 纹理强度估计的一种依据。

除此之外，自 H.264/AVC 起，新的视频编码标准均采用可变尺寸运动补偿技术，根据待编码 CU 的运动情况，可将其划分为多个不同尺寸的 PU 进行独立的运动估计和运动补偿，实现更加精细的预测编码。以 H.264/AVC 为例，当待编码 CU 以 16×16 的 PU 进行运动补偿时，表示该 CU 内的所有对象运动方向完全或趋于完全一致；反之，当待编码 CU 被划分为多个小尺寸 PU 进行运动补偿时，表示该 CU 内包含多个子对象，且其运动方向各不相同。

本文提出利用运动矢量和帧间预测分割模式对 P 帧编码块的纹理强度进行估计。

首先，根据每个 PU 的尺寸大小为其定义对应的权值因子：

$$w_{mn} = \lceil (\max\_h \max\_v)/(mn) \rceil \tag{9}$$

式中，w_{mn} 表示尺寸为 mn 的 PU 的权值因子；max_h 和 max_v 代表编码标准所支持的最大运动补偿编码块的尺寸。CU 划分越精细，表示 CU 内各个 PU 之间差异越大，越有可能成为纹理区域，所以对应 PU 权值应越高。以 H.264/AVC 视频编码为例，4×4 PU 的权值因子为 16，4×8 PU 的权值因子为 8，依次类推，16×16 PU 的权值因子为 1。

其次，提出一种尺寸为 mn 的 PU 的纹理强度量化方法，如式 (10) 所示：

$$\mathrm{TS}_{mn} = w_{mn}(\mathrm{mv}\_x + \mathrm{mv}\_y) \tag{10}$$

类似地，设定一个纹理强度阈值 T_P，当 TS_{mn} 大于或等于设定阈值 T_P 时，认为该区域为纹理区域，否则为非纹理区域，如式 (11) 所示：

$$\begin{cases} \mathrm{TS}_{mn} \geqslant T_P & \text{纹理区域} \\ \text{otherwise} & \text{非纹理区域} \end{cases} \tag{11}$$

正如 2.1 节中所述，阈值 T_P 同样应该随视频内容变化而变化。参考 2.1 节中的 TSSUM-I 模型，本文提出一种 P 帧阈值可伸缩自适应更新算法 (threshold scalable self-adaptive update method for p-frame, TSSUM-P)，其基本思想是：在同一个 GOP 中，根据前一帧的平均纹理强度和伸缩因子更新下一帧的纹理强度阈值，具体过程如下：

1) 为 GOP 中第 1 个待编码 P 帧设置初始阈值：

$$T_P(1) = T_{P_0} \tag{12}$$

2) 利用式 (10) 计算当前编码 P 帧 (定义为当前 GOP 中第 n 个 P 帧) 中 (x,y) 位置待编码块的纹理强度 $\mathrm{TS}_P(n,x,y)$ 并统计当前编码 P 帧中所有 PU 的纹理强度和：

$$T_{\mathrm{sum}-P} = \sum_i \sum_j (\mathrm{TS}_P(n,i,j)) \tag{13}$$

3) 根据当前编码 P 帧的平均纹理强度与伸缩

因子更新下一个待编码 P 帧 (定义为第 n+1 个 P 帧) 的纹理强度阈值：

$$T_P(n+1)=\lambda_P \text{mean}(T_{\text{sum}-P}) \tag{14}$$

式中，λ_P为伸缩因子，其值大于零。

4) 若第 n+1 个待编码 P 帧与第 n 个 P 帧不在同一个 GOP 内，则跳转到过程 1)，重置阈值；否则，跳转到过程 2) 继续执行。

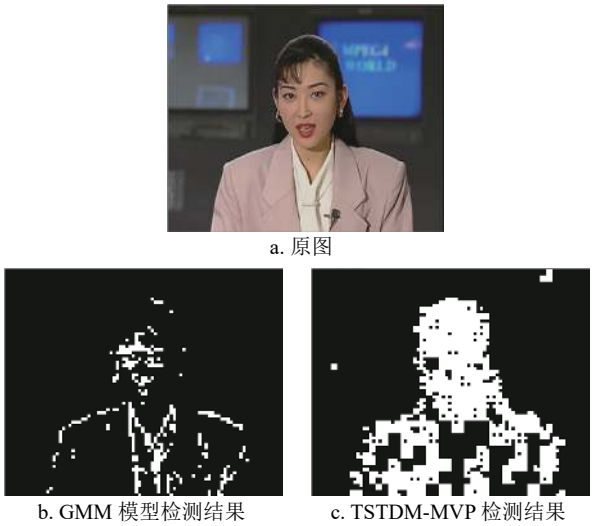
然而，视频编码协议中对 MV 采用了预测编码方式，由于存在空间相关性，基于相邻块运动矢量得到的当前编码块运动矢量预测值 (motion vector prediction, MVP)，可以很大程度上代表其真实值。为了保证解码端能够顺利提取纹理区域，对式 (10) 进行修正，利用 MVP 进行纹理估计：

$$\text{TS}_{mn}=w_{mn}(\text{mvp\_}x+\text{mvp\_}y) \tag{15}$$

式中，mvp_x 和 mvp_y 为 MVP 的两个分量。

至此，得到一个基于 MVP 的阈值自适应纹理检测模型 (threshold self-adaptive texture detection model based on MVP, STDM-MVP)。

同样，为了验证 TSTDM-MVP 纹理检测模型的有效性，本文以经典的 GMM 算法为参考模型，对两种算法的纹理检测结果进行对比，如图 5 所示。



a. 原图　b. GMM 模型检测结果　c. TSTDM-MVP 检测结果

图 5　纹理检测结果对比

从图 5 中可以看出，在对纹理检测精度要求不高的场景下，如选择性视频加密，本文提出的 TSTDM-MVP 纹理检测模型能够有效的检测出视频纹理区域；同时，根据式 (11) 和式 (15)，对于一个 16×16 的编码块，本文纹理检测算法平均需要 14 次加法和 1 次判断，而 GMM 算法计算过程复杂，折合加法运算约为 15 000 次，计算量远高于本文算法。

2.3　视频加密算法

为保证视频压缩编码效率和码流格式兼容性，根据前文中的分析，对于编码 I 帧，本文仅对编码块量化后的非零 DCT 系数符号进行加密；对于编码 P 帧，由于编码标准采用预测方式对运动矢量进行差分编码，故而选择 P 帧编码块非零 MVD 符号进行加密。

对于加密方式，本文采用实时性好的流加密算法。同时，为提高加密算法安全性，克服已知明文条件下的统计差分攻击，本文参考动态 S-Box 加密技术[25-26] 和分组加密 CFB 模式，提出一种新颖的融合视频特征的流密钥生成算法 (stream secret key generating method fusing video feature, SSKGM-VF)，用于流加密过程，如图 6 所示。

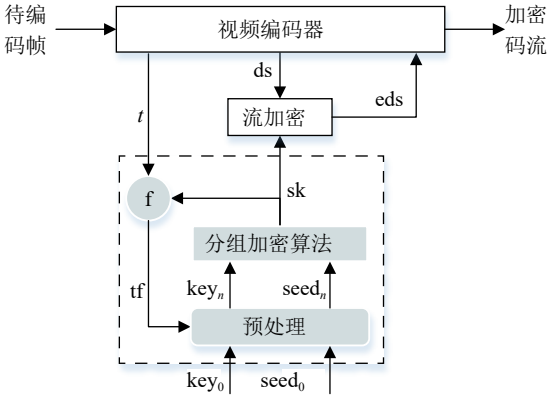


图 6　SSKGM-VF 算法结构图

详细的流密钥生成过程如下：

1) 视频编码开始前，根据用户输入的初始密钥 key$_0$，直接利用分组加密算法对初始种子 seed$_0$ 进行加密，得到一组加密数据，即图 6 中的 sk，用作流加密的初始流密钥；

2) 视频编码器利用前文中提出的纹理检测模型选出纹理块，输出纹理块的特征信息 t 并与 sk 进行 f 操作，得到的 tf 与 key$_0$ 和 seed$_0$ 一起输入到预处理模块进行处理，得到新的密钥 key$_n$ 和种子 seed$_n$，用于分组加密产生新的流密钥 sk；

3) 若视频编码尚未完成，返回过程 2) 继续执行；否则，流程结束。

利用 SSKGM-VF 算法生成的流密钥，对 I 帧编码块量化后的非零 DCT 系数和 P 帧编码块非零 MVD 的符号位进行加密，具体加密过程如式 (16) 所示：

$$\mathrm{ENCS} = \mathrm{NCS} \otimes \mathrm{sk}$$

$$\mathrm{EMVDS} = \mathrm{MVDS} \otimes \mathrm{sk} \tag{16}$$

式中，NCS 和 MVDS 分别代表纹理区域的非零 DCT 系数符号和非零 MVD 符号；ENCS 和 EMVDS 分别对应 NCS 和 MVDS 加密后的结果；sk 为加密密钥，由 SSKGM-VF 算法产生；⊗表示流加密运算。

2.4　视频解密算法

视频解密是视频加密的逆过程。编码端通过安全信道将初始密钥和初始种子传输到解码端，解码端按照如下步骤进行解密：

1) 解码器利用初始密钥 key_0 和初始种子 seed_0，计算流加密操作的初始流密钥 sk；

2) 解码器对视频压缩码流进行解析，得到加密后的量化 DCT 系数和 MVD，同时，利用预测机制计算得到 MVP；

3) 根据步骤 2) 得到的量化 DCT 系数和 MVP，解码器利用 TSTDM-QDCTC 和 TSTDM-MVP 分别对 I 帧和 P 帧进行纹理区域检测；

4) 对于纹理区域，利用流密钥 sk 对量化 DCT 系数或 MVD 进行符号解密；

5) 将正确解密后的视频特征反馈到 SSKGM-VF 算法中，继续产生新的流密钥用于解密过程；

6) 若压缩码流均已解密完成，则结束流程，否则返回步骤 2) 继续执行。

通过上述步骤，解码端即可得到真实的量化 DCT 系数和 MVD，后续视频解码过程才能得到真实的视频内容。

3　实验结果与分析

为验证本文所提加密算法的有效性，以应用广泛的 H.264/AVC 编解码器为验证平台、JM8.6 软件编解码器参考模型为实现平台，对 21 个不同场景、不同运动幅度的 CIF(352×288) 格式的标准视频序列进行测试，所有视频序列均按照帧率为 30 fps、I 帧编码周期为 30、I 帧量化参数 QP=28 的配置进行编码加密，总编码帧数为 150 帧。实验过程中，分组加密算法采用国密 SM4 实现，流加密运算、f 运算和预处理过程均采用逻辑异或代替实现。本文主要从视频加密效果、加密算法复杂度和加密算法安全性 3 个方面分析本文所提加密算法的性能。同时，为避免不同实验条件对实验结果的影响，测试过程中两种加密算法使用完全相同的流密钥进行加密。

3.1　视频加密效果

视频加密效果从主观感受和客观评价两个方面进行分析，且以视频加密前后的峰值信噪比 (peak signal noise ratio, PSNR) 作为客观评价的依据。本文以文献 [13] 提出的选择性加密算法为对照，对本文所提算法进行性能测试与评估。同时，为避免不同实验条件对实验结果的影响，测试过程中两种加密算法使用完全相同的加密密钥。图 7 为两个不同场景的测试视频在两种加密算法下的加密效果图。

a. 加密前 (从左至右为 bus、football)

b. 参考算法加密结果 (从左至右为 bus、football)

c. 本文算法加密结果 (从左至右为 bus、football)

图 7　加密效果对比

从图 7 中可以发现，主观上，与原始视频相比，本文所提加密算法与参考算法均能有效保护视频信息，人眼无法正确识别出加密视频的真实内容。不同的是，本文加密算法在平滑区域，如 bus 视频画面中的石碑建筑和画面深处相对较暗的树木、football 视频画面中的大部分草地，没有进行过多的加密，而是将其判定为非纹理区域进行处理；相反，本文算法在汽车、栏杆和运动员等细节较丰富的纹理区域、运动区域进行了有效加密。但是，参考算法[13] 没有考虑这一点，而是对全局视频进行同强度加密，虽然整体加密效果更混乱，但加密数据量较多。

客观上，本文对两种加密算法加密前后的视频信噪比进行了统计分析，如表 1 所示。

表 1 不同视频序列加密前后 PSNR 对比结果

dB

测试序列	加密前	文献[13]	本文
akiyo	48.44	9.96	12.77
bowing	46.93	8.74	10.86
bridge	54.67	8.54	11.43
bus	34.30	9.03	9.70
coastguard	33.99	10.98	10.09
container	41.27	10.66	8.33
flower	33.99	9.18	11.36
football	34.98	12.95	13.35
foreman	44.47	10.97	7.17
hall	40.89	12.11	9.92
highway	40.78	9.24	11.96
husky	24.19	9.58	11.16
ice	46.12	12.13	11.81
mobile	32.04	9.12	9.45
mother	45.15	10.74	12.89
news	45.40	10.37	12.34
paris	39.19	9.23	9.15
silent	42.02	9.81	8.89
soccer	40.36	13.08	12.06
tempete	34.59	9.15	10.19
waterfall	38.15	11.01	9.40
平均值	40.09	10.31	10.68

从表 1 中可以看出，对于 21 个不同场景的测试视频，按照本文所提的选择性加密算法进行加密，加密后视频平均信噪比为 10.68 dB，而利用文献 [13] 提出的算法进行加密，加密后视频平均信噪比为 10.31 dB，相比之下，本文算法平均高出 0.37 dB，这是因为本文算法只对纹理区域进行加密处理，而文献 [13] 则是进行全局加密，数据加密率高。但是，根据人眼视觉特性，当信噪比低于 15 dB 后，人眼将无法分辨出视频内容。因此本文加密算法和参考算法均能起到很好的视频内容加密作用。

3.2 加密算法复杂度

加密算法的复杂度主要取决于加密运算自身复杂度和加密数据量大小。与已有文献相同，本文加密运算依然采用逻辑异或实现，其本身计算复杂度极低，可不予考虑，因此，加密数据量成为衡量加密算法复杂度的唯一标准。

对于编码 I 帧的加密过程，除非零系数符号 (non-zero coefficient sign, NCS) 外，参考算法[13] 还对编码块的 IPM 进行了加密操作，具体加密数据量比较结果如图 8 所示。

对于 NCS 的加密过程，本文使用 TSTDM-QDCTC 纹理检测模型对加密区域进行了筛选，所以，相对于参考算法[13]，本文针对 NCS 的加密数据量明显减少，平均降低了 21.30%。然而，对于 IPM，视频编码协议对其有一定约束，即使加密后的 IPM 符合标准允许的若干种模式，但无法保证时间和空间上的相关规定，例如，加密后的 IPM 可能需要利用在时间或空间上位于当前编码块之后的相邻块数据进行预测，此类情况在通用解码器上无法正常解码。因此，本文算法并未对 IPM 进行加密，即针对 IPM 的加密数据量降低了 100%。

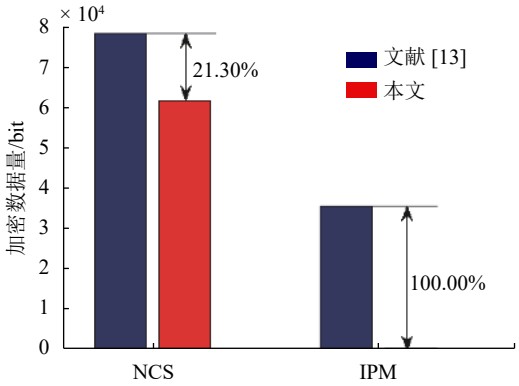


图 8 I 帧加密数据量对比

对于编码 P 帧的加密过程，本文算法和参考算法[13] 均选择 MVD 符号位进行加密，加密数据量比较结果如表 2 所示。

表 2 不同视频序列 MVD 符号加密数据量对比结果

测试序列	文献[13]/bit	本文/bit	减少量
akiyo	75 880	34 260	54.85%
bowing	142 504	55 704	60.91%
bridge	62 580	42 761	31.67%
bus	167 288	59 072	64.69%
coastguard	137 067	56 321	58.91%
container	42 247	24 028	43.12%
flower	123 822	42 550	65.64%
football	214 381	65 102	69.63%
foreman	170 630	60 148	64.75%
hall	84 396	32 499	61.49%
highway	96 142	52 039	45.87%
husky	183 931	43 978	76.09%
ice	189 515	40 585	78.58%
mobile	154 963	64 118	58.62%
mother	172 368	46 507	73.02%
news	118 743	45 581	61.61%
paris	122 376	34 474	71.83%
silent	126 740	41 020	67.63%
soccer	188 626	47 774	74.67%
tempete	169 899	44 511	73.80%
waterfall	93 582	57 447	38.61%
平均值	135 128	47 166	65.10%

从表 2 中可以发现，经过本文所提的 TSTDM-MVP 纹理检测模型选择后，本文算法对 MVD 符号的加密数据量大大减小，平均结果为 47 166 比特。而参考算法[13] 对编码 P 帧中的所有 MVD 的符号均进行了加密处理，平均加密数据量为 135 128 比特，相比之下，本文算法平均加密数据量减少了 65.10%，计算复杂度大大降低。

总体上，综合考虑所有加密对象后，本文加密算法相对于参考算法，加密数据量平均降低了 56.29%，若将加密数据量直接映射为加密时间开销，则本文加密算法时间消耗平均降低 56.29%，为实时性视频编码加密提供了有利条件。

3.3　加密算法安全性

首先，对加密算法密钥空间进行分析，本文加密算法主要有两类依赖：1) TSTDM-QDCTC 和 TSTDM-MVP 纹理检测模型；2) SSKGM-VF 流密钥生成算法。其中，纹理检测模型共有 4 个参数：T_{I0}、λ_I、T_{P0} 和 λ_P；流密钥生成算法则有 3 个参数：key_0、seed_0 和 t。因此，本文加密算法密钥空间 K 可表示为如下形式：

$$K=\{T_{I0},\lambda_I,T_{P0},\lambda_P,\text{key}_0,\text{seed}_0,t\} \tag{17}$$

式中，key_0、seed_0 和 t 均为 128 bit，在仅考虑 key_0 和 seed_0 的情况下，密钥空间为 $2^{128}\times2^{128}=2^{256}$，有效密钥长度已达 256 bit，而一般认为算法密钥长度达到 128 bit 即为安全的，因此，本文加密算法密钥空间达到安全标准，能够有效抵抗暴力攻击。

此外，本文利用 NIST SP 800-22 对 SSKGM-VF 算法生成的流密钥随机性进行了测试。NIST SP 800-22 是由美国国家标准与技术研究院公布的数据序列随机性测试工具，测试结果由 P 值决定，当 P 值大于 0.01 时，认为被测试序列是随机的。详细测试结果如表 3 所示。

表 3　流密钥 SP 800-22 测试结果

测试项	P值	结果
Frequency	0.455 937	通过
Block frequency	0.494 392	通过
Cumulative Sums	0.964 295	通过
Runs	0.935 716	通过
Longest Run	0.798 139	通过
Rank	0.798 139	通过
FFT	0.739 918	通过
Non-overlapping template	0.867 692	通过
Approximate entropy	0.383 827	通过
Serial	0.964 295	通过
Linear complexity	0.401 199	通过

从表 3 中可以得出结论，基于 SSKGM-VF 算法生成的流密钥具有较好的随机性，能够抵御静态分析攻击。

4　结 束 语

本文提出了一种基于视频混合编码器的低复杂度纹理自适应视频加密算法，通过利用编码过程中间数据 (残差 DCT 系数和运动矢量)，构建了两种低计算复杂度的纹理检测模型 TSTDM-QDCTC 和 TSTDM-MVP，并以此作为加密区域选择算法，实现对纹理区域的有效加密。以 H.264/AVC 视频编码为测试平台进行了算法验证，实验结果表明，本文提出的纹理检测模型和密钥生成算法，加密后的视频内容得到了很好的保护，平均信噪比为 10.68 dB；与此同时，相比于已有的 SE 算法，加密数据量大大减小，平均降低了 56.29%，为实时性视频编码加密提供了有利条件。

参　考　文　献

[1] 程东升, 谭旭, 许志良, 等. 结合思维超混沌系统和位分解的图像加密算法研究[J]. 电子科技大学学报, 2018, 47(6): 906-912.
CHENG Dong-sheng, TAN Xu, XU Zhi-liang, et al. Image encryption algorithm research by combining four dimensional hyper-chaotic system and bit decomposition[J]. Journal of University of Electronic Science and Technology of China, 2018, 47(6): 906-912.

[2] ZENG W, LAN J, ZHUANG X. Security for multimedia adaptation: architectures and solutions[J]. IEEE Multimedia, 2006, 13(2): 68-76.

[3] MASSOUDI A, LEFEBVRE F, DE VLEESCHOUWER C, et al. Overview on selective encryption of image and video: Challenges and perspectives[J]. EURASIP Journal on Information Security, 2008, 2008(1): 1-18.

[4] WANG L F, WANG W D, JIAN M A, et al. Perceptual video encryption scheme for mobile application based on H.264[J]. The Journal of China Universities of Posts and Telecommunications, 2008, 15(S): 73-78.

[5] AHN J, SHIM H J, JEON B, et al. Digital video scrambling method using intra prediction mode[C]//Pacific-Rim Conference on Multimedia. Berlin: Springer-Verlag, 2004: 386-393.

[6] LI Y, LIANG L, SU Z, et al. A new video encryption algorithm for H.264[C]//International Conference on Information. [S.l.]: IEEE, 2005: 1121-1124.

[7] KHLIF N, DAMAK T, KAMMOUN F, et al. A very efficient encryption scheme for the H.264/AVC CODEC adopted in intra prediction mode[C]//Image Processing, Applications & Systems Conference. [S.l.]: IEEE, 2014: 1-7.

[8] WANG Y, O'NEILL M, KURUGOLLU F. The improved sign bit encryption of motion vectors for H.264/AVC[C]// Signal Processing Conference. [S.l.]: IEEE, 2012: 1752-

1756.
[9] KHLIF N, DAMAK T, KAMMOUN F, et al. Motion vectors signs encryption for H.264/AVC[C]//International Conference on Advanced Technologies for Signal & Image Processing. [S.l.]: IEEE, 2014: 1-6.
[10] RAJAGOPAL S, SHENBAGAVALLI A. Design of real time video encryption system based on adaptive elastic motion model in H.264[J]. Journal of Computational and Theoretical Nanoscience, 2016, 13(10): 7156-7170.
[11] MAO N, ZHUO L, ZHANG J, et al. Fast compression domain video encryption scheme for H.264/AVC streaming[C]//International Conference on Advanced Communication Technology. [S.l.]: IEEE, 2012: 425-429.
[12] TEW Y, MINEMURA K, WONG K S. HEVC selective encryption using transform skip signal and sign bin[C]// 2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA). [S.l.]: IEEE, 2015: 963-970.
[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, 2016, 76(3): 3235-3253.
[14] XU H, TONG X J, ZHANG M, et al. Dynamic video encryption algorithm for H264/AVC based on a spatiotemporal chaos system[J]. Journal of the Optical Society of America A, 2016, 33(6): 1166.
[15] DI X, WANG Y, LI J, et al. An optimized video selective encryption algorithm[C]//2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). [S.l.]: IEEE, 2017: 1-5.
[16] CHEN J, PENG F, LONG M. A perceptual encryption scheme for HEVC video with lossless compression[C]// International Conference on Cloud Computing and Security. Cham: Springer, 2017: 396-407.
[17] SALLAM A I, FARAGALLAH O S, EL-RABIE E S M. HEVC selective encryption using RC6 block cipher technique[J]. IEEE Transactions on Multimedia, 2017, 20(7): 1636-1644.
[18] 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.
[19] DUFAUX F, EBRAHIMI T. Scrambling for privacy protection in video surveillance systems[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 8(18): 1168-1174.
[20] PENG F, ZHU X W, LONG M. An ROI privacy protection scheme for H.264 video based on FMO and Chaos[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(10): 1688-1699.
[21] SIDATY N, VIITANEN M, HAMIDOUCHE W, et al. Live demonstration: end-to-end real-time ROI-based encryption in HEVC videos[C]//IEEE International Symposium on Circuits & Systems. [S.l.]: IEEE, 2018: 1-1.
[22] WANG Y, O"NEILL M, KURUGOLLU F. A tunable encryption scheme and analysis of fast selective encryption for CAVLC and CABAC in H.264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(9): 1476-1490.
[23] ZHAO Y, SHEN H, ZHUO L. An Efficient motion reference structure based selective encryption algorithm for H.264 videos[J]. IET Information Security, 2014, 8(3): 199-206.
[24] LOÏC DUBOIS, PUECH W, BLANC-TALON J. Smart selective encryption of H.264/AVC videos using confidentiality metrics[J]. Annals of Telecommunications-Annales Des Télécommunications, 2014, 69(11-12): 569-583.
[25] HONG S S, HAN M M. The study of selective encryption of motion vector based on the s-box for the security improvement in the process of video[J]. Multimedia Tools and Applications, 2014, 71(3): 1577-1597.
[26] ALTAF M, AHMAD A, KHAN F A, et al. Computationally efficient selective video encryption with chaos based block cipher[J]. Multimedia Tools and Applications, 2018, 77: 27981-27995.

编 辑 刘飞阳