DOI:10.11918/202212091

色度域亮度域信息融合的监控视频重压缩取证

公衍超[1,2,3],王子琳[1],杨楷芳[4],刘 颖[1,2,3],林庆帆[1,5],王富平[1,2,3]

(1.西安邮电大学 通信与信息工程学院,西安 710121;
2.西安邮电大学 陕西省法庭科学电子信息实验研究中心,西安 710121;
3.西安邮电大学 陕西省无线通信与信息处理技术国际联合研究中心,西安 710121;
4.陕西师范大学 计算机科学学院,西安 710119; 5.新加坡 XsecPro 公司,新加坡 787820)

摘 要:为解决当前视频重压缩取证方法没有考虑色度域信息、取证准确度低的问题,提出一种面向最新多用途视频编码(versatile video coding,VVC)标准色度域亮度域信息融合的监控视频重压缩取证方法(CLF-SVRF)。基于VVC标准的编码原理,从监控视频的色度域和亮度域维度分析并确定VVC视频码流中与压缩次数密切相关的基础码流特征;基础码流特征包括色度域和亮度域编码单元(coding unit,CU)的划分类型及预测模式;结合拉格朗日率失真优化技术分析随着压缩次数的增加,色度域亮度域CU划分类型和预测模式的变化;进一步确定色度域亮度域CU划分类型和预测模式可以作为检测视频压缩次数的基础码流特征;接着考虑视频监控应用对重压缩取证方法低复杂度的需求,基于色度域亮度域CU划分类型和预测模式构建低复杂度高级码流特征;将高级码流特征输入支持向量机完成监控视频的重压缩取证。实验结果表明,与当前先进方法相比,CLF-SVRF方法的监控视频重压缩取证准确度平均提升了13.53%,同时可以大幅度地降低重压缩取证耗时,重压缩取证时间平均减少了47.42%。

关键词:视频编码;通用视频编码;重压缩取证;监控视频;色度域
中图分类号:TN911.7 **文献标志码**:A **文章编号**:0367-6234(2024)05-0046-10

Video forensic compression with chroma and luminance domain information fusion for surveillance videos

GONG Yanchao[1,2,3], WANG Zilin[1], YANG Kaifang[4], LIU Ying[1,2,3], LIM Kengpang[1,5], WANG Fuping[1,2,3]

(1. School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China;
2. Electronic Information Experimental Research Center for Forensic Science of Shaanxi Province, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; 3. International Joint Research Center for Wireless Communication and Information Processing Technology of Shaanxi Province, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; 4. School of Computer Science, Shaanxi Normal University, Xi'an 710119, China;
5. XsecPro Pte. Ltd., Singapore 787820, Singapore)

Abstract: To address the issues of not considering chromatic domain information and low accuracy in current video froensic compression methods, a surveillance video recompression forensics method for the latest versatile video coding (VVC) standard based on the information fusion of chroma domain and luminance domain, referred to as CLF-SVRF, is proposed. Based on the coding principles of the VVC standard, the basic bitstream characteristics in VVC video bitstream that are closely related to compression time are analyzed and determined from the dimensions of chroma domain and luminance domain in surveillance videos. The basic bitstream characteristics include the partitioning type and prediction mode of coding unit (CU) in chroma domain and luminance domain. Combining with the Lagrangian rate distortion optimization technique, the variations in the partitioning type and prediction mode of CU in chroma domain and luminance domain as the compression time increases are analyzed. It is further determined that the partitioning type and prediction mode of CU in chroma domain and luminance domain can be used as basic bitstream characteristics for detecting the video compression time. Then, considering the requirement of video surveillance applications for low complexity forensics methods, a low complexity advanced bitstream characteristics is constructed based on the partitioning type and prediction mode of CU in chroma domain and luminance domain. The advanced bitstream characteristics are input into the support vector machine to complete the

收稿日期:2022-12-29;**录用日期**:2023-04-11;**网络首发日期**:2023-04-27
网络首发地址:http://kns.cnki.net/kcms/detail/23.1235.T.20230427.1149.008.html
基金项目:国家自然科学基金(62277036)
作者简介:公衍超(1986—),男,副教授,硕士生导师
通信作者:公衍超,gongyanchao@xupt.edu.cn

recompression forensics of surveillance videos. The experimental results show that compared with the current advanced methods, the CLF-SVRF method can improve the accuracy of surveillance video recompression forensics by 13.53% on average. At the same time, it can significantly reduce the time required for forensic recompression, and reduce the recompression forensics time by 47.42% on average.

Keywords: video coding; versatile video coding; recompression forensics; surveillance video; chroma domain

当前视频监控系统已经深度参与到刑事侦查、司法取证、防恐反恐等领域，为社会的安全稳定提供了关键保障。但是近年来，随着图像编辑软件功能的日益强大和深度篡改技术的飞速发展，对监控视频进行篡改越来越容易。保障监控视频数据安全性已经成为社会的普遍共识。相应地，检测视频内容是否被篡改的视频取证技术被广泛关注[1-2]。

在视频监控系统中，摄像头采集到的原始监控视频只需要编码压缩一次就通过网络传输给用户。编码压缩后的视频码流在网络传输时及在用户端存储时都有可能被篡改者盗取。篡改者将盗取的视频码流解码并篡改视频内容，然后必须将篡改视频再次编码压缩后才能得到包含篡改信息的视频码流从而完成整个篡改操作。所以编码压缩次数成为判定监控视频是否经过篡改的重要依据，以检测视频码流编码压缩次数为目的的视频重压缩取证技术应运而生。

视频重压缩取证方法的一个关键是提取视频码流中与压缩次数密切相关的特征。所以视频重压缩取证方法与具体的视频编码标准密切相关。当前在视频监控系统中被广泛应用的国际视频编码标准是 2003 年发布的 H.264/AVC 标准[3]和 2013 年发布的 H.265/HEVC 标准[4]。文献[5-9]面向 H.264/AVC 分别提出了有效的视频重压缩取证方法，考虑的视频码流特征包括块效应强度、量化信息、离散余弦变换系数、宏块类型、运动矢量信息等。文献[10-19]面向 H.265/HEVC 提出了有效的视频重压缩取证方法，考虑的视频码流特征包括变换单元类型、预测单元模式、环路滤波特征、图像组结构信息等。文献[20]面向 2020 年发布的最新一代国际视频编码标准 H.266/VVC[21]提出了考虑亮度域编码单元(coding unit，CU)预测模式和分区模式的重压缩取证方法。当前提出的重压缩取证方法还存在以下不足：

1)大部分方法都是面向早期的视频编码标准，不适用于最新的 H.266/VVC 标准。相比于之前标准，H.266/VVC 采用了更加先进的编码技术，在显著提高编码压缩效率的同时，也改变了视频码流的语法语义与基础特征。

2)在构建重压缩特征时只考虑了亮度域编码信息，忽视了色度域编码信息。色度域信息反映了监控视频中颜色内容变化，对于理解监控视频内容，例如嫌疑人着装、嫌疑车辆涂装信息等，具有重要影响。H.266/VVC 中也采用了很多专门针对色度域信息的编码技术。这导致相比于亮度域，色度域会提供一些不同的编码信息，可以有效提升重压缩取证效率。

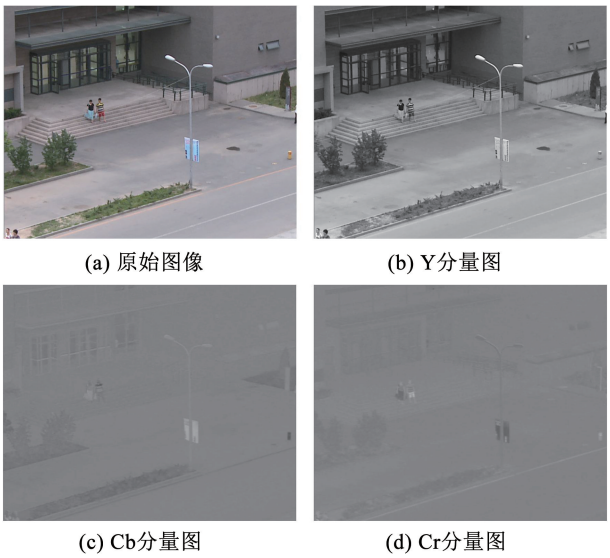
3)复杂度较高，不适于在视频监控系统应用。例如文献[8-9,16-17,20]提出的方法需要对待取证的视频码流重新编码压缩超过 1 次，这显著增加了取证复杂度。视频监控系统是典型的 7 d+24 h 全天候工作系统，与案件相关的视频片段动辄达到百、千的级别。大数量视频片段和案件破获时效性决定了视频监控应用对视频重压缩取证方法有低复杂度的需求。

针对以上不足，本文提出了面向 H.266/VVC 标准，结合考虑色度域与亮度域编码信息的低复杂度重压缩取证方法。

1　H.266/VVC 相关编码原理

1.1　YCbCr 颜色空间

视频编码通常压缩的是 YCbCr 颜色空间视频。Y 是亮度分量，Cb 和 Cr 是色度分量，分别表征亮度域和色度域上的视频信息。图 1 所示为标准监控视频 Classover 序列第 1 帧的原始图像及 3 个分量图 Y、Cb、Cr。Y 分量图对应的是原始图像的灰度图，反映像素亮度值的变化。在 Y 分量基础上添加 Cb 和 Cr 分量信息即可得到原始彩色图像。



(a) 原始图像　(b) Y分量图　(c) Cb分量图　(d) Cr分量图

图 1　Classover 序列的第 1 帧

Fig. 1　First frame of Classover sequence

编码视频采用 YCbCr 颜色空间的主要原因是人眼对视频亮度分量和色度分量的感知敏感度不同[22]。采用 YCbCr 颜色空间易在编码时将亮度分量和色度分量信号分开用不同的技术处理以提升各自的编码效率。亮度分量和色度分量不同的内容特性及采用的不同编码技术最终导致了其不同的编码压缩特性。

1.2 编码单元划分类型

H. 266/VVC 采用基于块的编码框架。根据是否使用帧间预测技术,将视频中的帧分为以下几种类型:只使用帧内预测技术不使用帧间预测技术的 I 帧和同时使用帧内预测技术和帧间技术的 P 帧与 B 帧。然后,将视频帧划分为编码树状单元(coding tree unit,CTU),再将 CTU 划分为 CU。H. 266/VVC 摒弃了 H. 265/HEVC 中预测单元和变换单元的概念,而是统一为 CU。同时,为了适应更加丰富的视频内容,H. 266/VVC 新采用了融合使用二叉树、三叉树和四叉树的多类型树划分技术以支持更灵活的 CU 划分方式,并且在帧内预测(intra prediction)帧中支持长宽不等的矩形 CU。考虑视频色度和亮度分量具有不同的内容特性,H. 266/VVC 新采用了色度分离树技术支持 I 帧中的色度和亮度分量独立地进行 CU 划分类型选择,支持的 CU 划分类型分别见表 1、2。

表 1 H. 266/VVC 支持的 I 帧色度分量 CU 划分类型

Tab. 1 I frame chroma component CU partitioning type supported by H. 266/VVC

编号	尺寸/像素	编号	尺寸/像素	编号	尺寸/像素
1	8×8	8	16×32	15	64×4
2	8×16	9	16×64	16	64×8
3	8×32	10	32×4	17	64×16
4	8×64	11	32×8	18	64×32
5	16×4	12	32×16	19	64×64
6	16×8	13	32×32		
7	16×16	14	32×64		

表 2 H. 266/VVC 支持的 I 帧亮度分量 CU 划分类型

Tab. 2 I frame luminance component CU partitioning type supported by H. 266/VVC

编号	尺寸/像素	编号	尺寸/像素	编号	尺寸/像素
1	4×4	7	8×16	13	32×4
2	4×8	8	8×32	14	32×8
3	4×16	9	16×4	15	32×16
4	4×32	10	16×8	16	32×32
5	8×4	11	16×16	17	64×64
6	8×8	12	16×32		

1.3 编码单元预测模式

预测模式是 H. 266/VVC 预测模块中 CU 需要选择的关键信息,其决定了预测值的来源。为了适应丰富的视频类型及纹理,H. 266/VVC 支持了更多种类的帧内预测模式。图 2 所示为 H. 266/VVC 亮度分量支持的帧内预测模式。

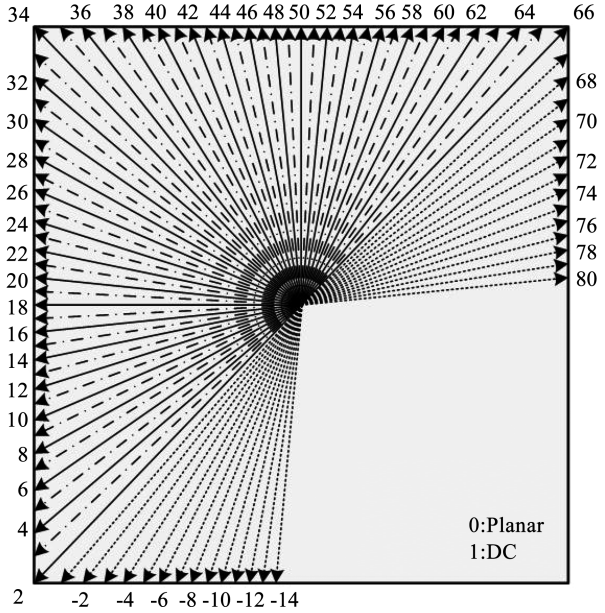


图 2 H. 266/VVC 亮度分量支持的帧内预测模式

Fig. 2 Intra prediction mode supported by H. 266/VVC luminance component

图 2 中实线表示 H. 265/HEVC 支持的 33 种角度预测模式。在此基础上 H. 266/VVC 将角度模式由 33 种扩展到了 65 种,新增加的模式由点划线表示。H. 266/VVC 同时保留了 Planar 模式和 DC 模式,并且为了提高非正方形 CU 的预测效率,H. 266/VVC 新采用了宽角度帧内预测技术[21],扩展了更多的预测方向,见图 2 中虚线。对于色度分量,除了角度模式、Planar 模式和 DC 模式外,H. 266/VVC 还增加了有效考虑色度分量与亮度分量之间相关性的跨分量线性模型(cross-component linear model,CCLM)模式[21]。

2 重压缩特征构建

2.1 基础码流特征确定

2.1.1 色亮度域 CU 划分类型与压缩次数关系

H. 266/VVC 采用拉格朗日率失真优化方法[21,23]从支持的 CU 划分类型中给每一个 CTU 选择一组最优的 CU 划分类型。最优的 CU 划分类型对应最小的拉格朗日率失真代价 J,具体表示为

$$\min J,\ J = \sum D_{i,j} + \hat{\lambda} \sum (\bar{R}_{i,j} + \hat{R}_{i,j}) \quad (1)$$

式中:$D_{i,j}$、$\bar{R}_{i,j}$、$\hat{R}_{i,j}$分别为 CTU 中第 i 个 CU 选择第 j

种 CU 划分类型产生的编码失真、纹理比特和头比特。编码码率由头比特和纹理比特两部分构成，其中纹理比特记录视频内容信息，头比特记录视频格式及编码参数信息。$\hat{\lambda}$ 为用于 CU 划分类型选择的拉格朗日因子，其取值主要与量化步长和图像类型相关，与 CU 划分类型无关，故 J 的取值主要与 $D_{i,j}$、$\bar{R}_{i,j}$ 和 $\hat{R}_{i,j}$ 相关。下面结合 H.266/VVC 编码原理，分析随着压缩次数的增加，大尺寸 CU 划分类型和小尺寸 CU 划分类型对应的 $D_{i,j}$、$\bar{R}_{i,j}$ 和 $\hat{R}_{i,j}$ 的取值变化，从而得出 CU 划分类型的选择规律。

H.266/VVC 采用基于预测、变换、量化、熵编码的混合帧内编码技术。预测是通过 CU 划分类型选择、CU 预测模式选择等子技术给当前编码 CTU 中的原始像素寻找最相关的预测像素，然后定义它们的差值为残差值并将其作为变换的输入。变换是将像素域的残差值转换为频域的变换系数，其中低频系数对应视频的结构信息，而高频系数对应视频的细节信息。通常变换系数具有较大的动态范围，需要对其进行量化，以减小变换系数的取值空间，提高熵编码去除熵冗余的能力，最终保证视频编码的效率。在上述过程中，一方面，H.266/VVC 采用多对一的量化映射机制，在减少变换系数取值空间的同时也会显著降低视频细节信息的表达能力，导致视频清晰度降低，视频内容变模糊；另一方面，考虑人类视觉系统对于视频细节信息感知不敏感，H.266/VVC 进一步采用了量化矩阵技术对高频系数使用更大的量化步长，这易导致高频信息的丢失，从而使得视频内容进一步变模糊。综上，原始监控视频随着压缩次数的增加，其视频内容清晰度会逐渐降低，视频内容会逐渐模糊。

模糊的视频内容对应模糊的 CTU，其包含的色度域或亮度域信息间的空域相关性增强。此时选择大尺寸 CU 划分类型或者小尺寸 CU 划分类型都能给原始像素找到相关性强的预测像素，对应残差值都很小，残差值分布近似趋近于零均值的拉普拉斯分布。残差值是后续变换量化熵编码的输入，直接决定了 $D_{i,j}$ 和 $\bar{R}_{i,j}$ 的大小，小的残差值及零均值拉普拉斯分布特性意味着小的 $D_{i,j}$ 和 $\bar{R}_{i,j}$。但是相比于大尺寸 CU 划分类型，小尺寸 CU 划分类型具有更加复杂的 CTU 多类型树划分数据，这些数据需要记录到头比特中，导致小尺寸 CU 划分类型对应的 $\hat{R}_{i,j}$ 更大。综上，随着视频压缩次数的增加，相比于大尺寸 CU 划分类型，小尺寸 CU 划分类型对应的 $D_{i,j}$、$\bar{R}_{i,j}$ 取值接近，但是 $\hat{R}_{i,j}$ 更大，所以 J 更大，其所占比例会逐渐降低。下面通过实验验证上述理论分析。

将表 1 和表 2 所示的 CU 划分类型按包含的像素数划分为小尺寸 CU 和大尺寸 CU 两类。包含像素数少于等于 256 的 CU 划分类型属于小尺寸 CU，其他的属于大尺寸 CU。用 H.266/VVC 参考软件 VTM11.0[24] 编码 Classover 序列 3 次。档次为 main_10，其他编码参数参照配置文件 encoder_lowdelay_P_vtm.cfg 中的默认设置。图 3 所示为 Classover 序列所有 I 帧中色度域和亮度域大尺寸及小尺寸 CU 占比随压缩次数的变化。

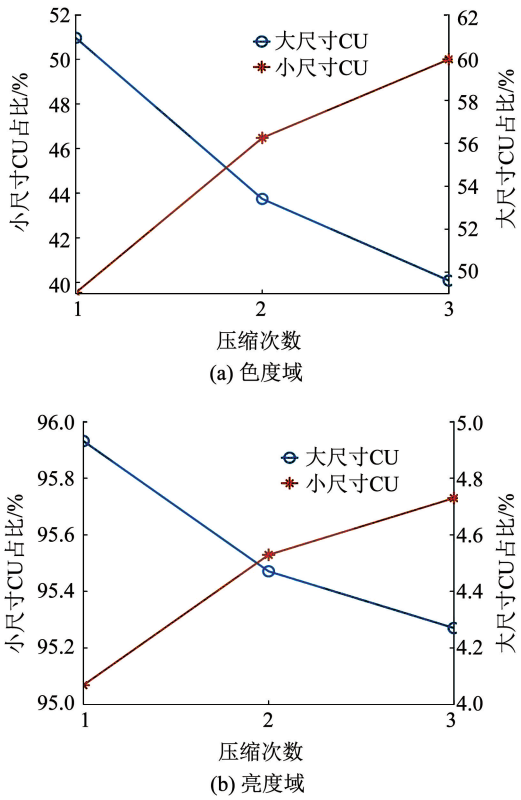
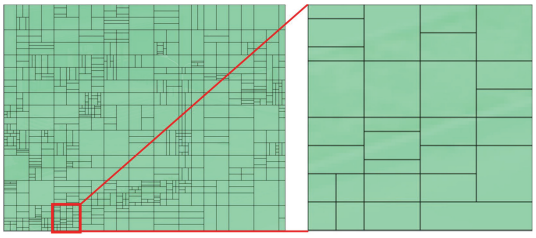


图 3　色度域亮度域 CU 划分类型占比随压缩次数的变化

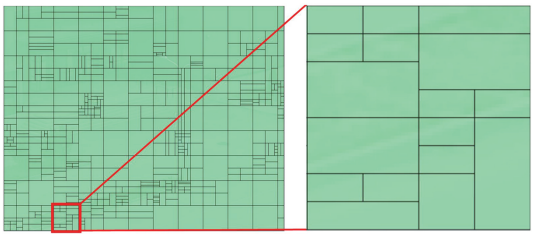
Fig. 3　Variation in CU partitioning type proportion of the chroma and luminance domain with compression times

由图 3 看出，随着压缩次数的增加，色度域和亮度域大尺寸 CU 占比逐渐增加，而小尺寸 CU 占比逐渐降低，并且前 2 次压缩对应的 CU 划分类型占比变化要比后 2 次压缩对应的 CU 划分类型占比变化更加剧烈。综上可以得到以下结论：视频压缩次数影响色度域亮度域 CU 划分类型的选择，且压缩次数越靠前影响越大。为了直观地展示上述结论，图 4、5 分别给出了 Classover 序列第 1 帧图像随着压缩次数增加其中色度域和亮度域 CU 划分类型的变化。图中不同尺寸的块表示不同的 CU 划分类型。由图中数据易看出，随着压缩次数增加，大尺寸 CU 在增加而小尺寸 CU 在减少。为了更加直观地

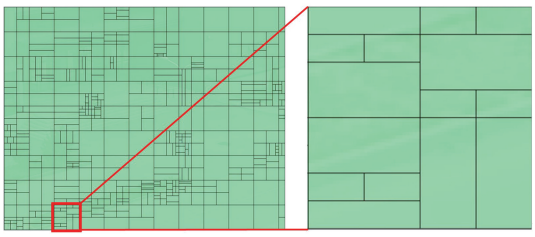
观察这一现象,也将左图中框线所示区域放大如右图展示。



(a) 第1次压缩图像的CU划分类型



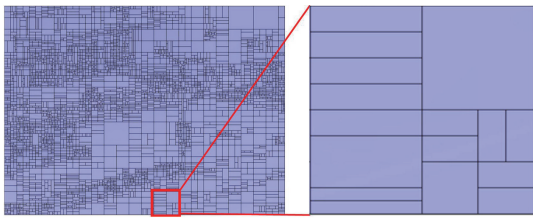
(b) 第2次压缩图像的CU划分类型



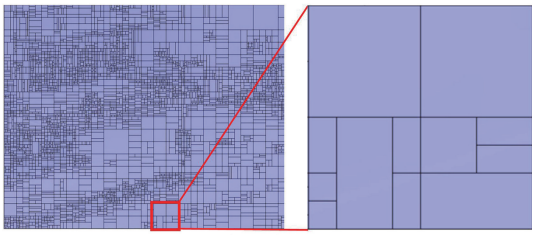
(c) 第3次压缩图像的CU划分类型

图 4 色度域 CU 划分类型随压缩次数的变化

Fig. 4 Variation in the chroma domain CU partitioning type with compression times



(a) 第1次压缩图像的CU划分类型



(b) 第2次压缩图像的CU划分类型



(c) 第3次压缩图像的CU划分类型

图 5 亮度域 CU 划分类型随压缩次数的变化

Fig. 5 Variation in the luminance domain CU partitioning type with compression times

2.1.2 色亮域 CU 预测模式与压缩次数关系

H.266/VVC 也是采用拉格朗日率失真优化方法给每一个 CU 选择最优的预测模式[21]。最优的预测模式对应最小的拉格朗日率失真代价 J,具体表示为

$$\min J, J = D_g + \bar{\lambda}(\bar{R}_g + \hat{R}_g) \tag{2}$$

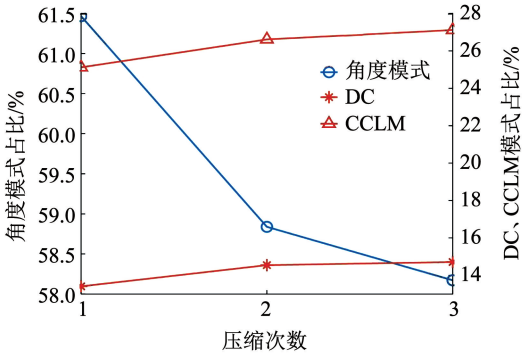
式中:D_g、$\bar{R}_g$、$\hat{R}_g$ 分别为当前编码 CU 选择第 g 种预测模式时产生的编码失真、纹理比特和头比特;$\bar{\lambda}$ 为用于 CU 预测模式选择的拉格朗日因子,其取值主要与量化步长和图像类型相关,与 CU 预测模式无关。

如图 2 所示,任一角度模式是利用特定一个方向的参考像素给当前编码像素确定预测值。Planar 模式则是利用当前编码像素正上方和平行左方的参考像素获得预测值,所以 Planar 模式可以认为是一种利用了 2 个角度信息的特殊角度模式。DC 模式是利用当前编码像素上方及左方的所有参考像素的平均值作为预测值。而 CCLM 模式是利用了色度与亮度分量之前的相关性确定预测值。前文已经从编码原理角度分析得到了原始监控视频随着压缩次数的增加,其视频内容清晰度会逐渐降低,视频内容会逐渐模糊。模糊的视频内容即意味着纹理方向信息的丢失,视频内容变的平坦,而相应的色度与亮度分量之间的相关性也会越强。所以,相比角度模式,DC 模式和 CCLM 模式更易找到优化的预测值。优化的预测值对应着更小的残差值,即对应着更小的 D_g 和 $\bar{R}_g$。而选择角度模式、DC 模式或者 CCLM 模式主要影响了头比特语法元素的取值,对于 $\hat{R}_g$ 影响不大。综上,随着压缩次数增加,角度模式(包括 Planar 模式)占比会降低,而 DC 和 CCLM 模式占比会增加。下面通过实验验证上述分析。

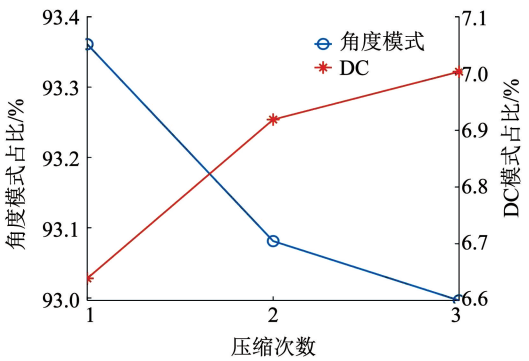
使用 VTM11.0 编码 Classover 序列 3 次,其他编码参数设置与前文实验设置一致。图 6 所示为 Classover 序列所有 I 帧中色度域和亮度域 CU 预测模式占比的变化。

由图 6 中数据易看出,随着视频压缩次数的增加角度模式(包括 Planar 模式)的占比降低,而 DC 和 CCLM 模式的占比增加。并且前 2 次压缩对应的预测模式占比变化要比后 2 次压缩对应的预测模式占比变化更加剧烈。综上可以得到以下结论:视频压缩次数影响色度域亮度域 CU 预测模式的选择,且压缩次数越靠前影响越大。为了更加直观地展示上述结论,图 7、8 分别给出了 Classover 序列第 1 帧图像随着压缩次数的增加其中色度域和亮度域 CU 帧内预测模式的变化。图中不同深度的色块表示不

同的 CU 帧内预测模式。由图中数据易看出压缩次数对于 CU 帧内预测模式的选择具有显著影响。



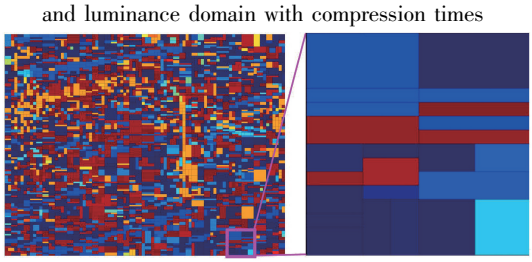
(a) 色度域



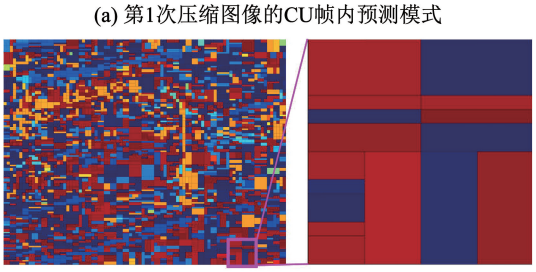
(b) 亮度域

图 6　色度域亮度域 CU 预测模式占比随压缩次数的变化

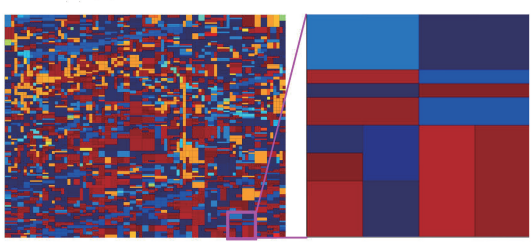
Fig. 6　Variation in CU prediction mode proportion of chroma and luminance domain with compression times



(a) 第1次压缩图像的CU帧内预测模式



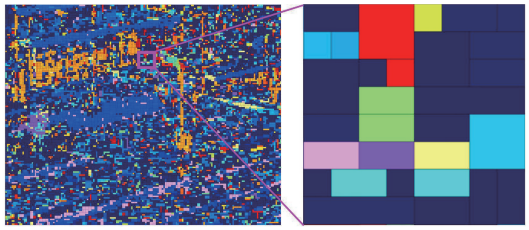
(b) 第2次压缩图像的CU帧内预测模式



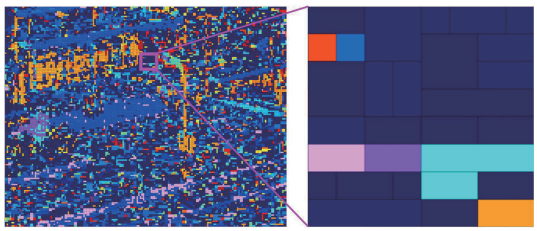
(c) 第3次压缩图像的CU帧内预测模式

图 7　色度域 CU 预测模式随压缩次数的变化

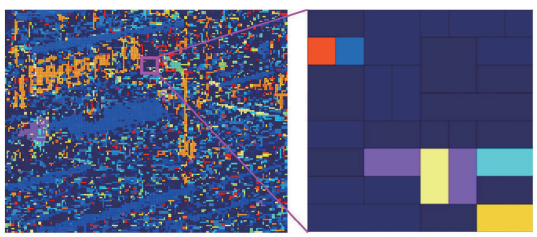
Fig. 7　Variation in chroma domain CU prediction mode with compression times



(a) 第1次压缩图像的CU帧内预测模式



(b) 第2次压缩图像的CU帧内预测模式



(c) 第3次压缩图像的CU帧内预测模式

图 8　亮度域 CU 预测模式随压缩次数的变化

Fig. 8　Variation in luminance domain CU prediction mode with compression times

2.2　高级码流特征构建

在构建高级码流特征时考虑了以下因素：1）色度域亮度域 CU 划分类型和预测模式与视频压缩次数具有明显关系，并且其会被写入到 H. 266/VVC 视频码流的头信息中，所以 CU 划分类型和预测模式可以作为检测视频压缩次数的基础码流特征；2）压缩次数越靠前，压缩次数对于 CU 划分类型和预测模式的影响越大。且在实际取证中考虑的压缩次数越多，高级码流特征计算通常会越复杂。当前一些重压缩取证代表性方法在进行本部分涉及的高级码流特征构建时考虑的是亮度域多次压缩信息[16]。而本文只考虑前 2 次压缩对应的色度域和亮度域 CU 划分类型和预测模式构建高级码流特征。特别说明，这里的前2 次指的是下文中的第 n 次和第 $n+1$ 次。

定义$C_{n,i,x}$为视频经过第 n 次压缩后第 i 个 I 帧第 x 个色度域 CU 包含划分类型和预测模式的信息矩阵，具体表示为

$$\boldsymbol{C}_{n,i,x}=[a_{n,i,x},b_{n,i,x}] \tag{3}$$

式中 $a_{n,i,x}$、$b_{n,i,x}$分别为第 n 次压缩视频第 i 个 I 帧第 x 个色度域 CU 选择的划分类型及预测模式编号。它们的取值范围为

$$a_{n,i,x}\in\{1,2,\cdots,\varphi\} \tag{4}$$

$$b_{n,i,x}\in\{1,2,\cdots,\varepsilon\} \tag{5}$$

式中 φ 和 ε 分别为 H.266/VVC 标准支持的色度域 CU 可选的划分类型和预测模式总数。

视频经过第 $n+1$ 次压缩后第 i 个 I 帧第 x 个色度域 CU 对应信息矩阵 $\boldsymbol{C}_{n+1,i,x}$ 的定义与 $\boldsymbol{C}_{n,i,x}$ 的定义类似，将式(3)～(5)中的 n 替换为 $n+1$ 即可，不再赘述。式(6)为色度域相同信息 CU 判别公式，具体表示为

$$\boldsymbol{C}_{n,i,x}=\boldsymbol{C}_{n+1,i,x} \tag{6}$$

对于第 n 次压缩视频第 i 个 I 帧第 x 个色度域 CU，假如其满足式(6)，则判定其为相同信息 CU，否则为不同信息 CU。

然后根据式(7)统计第 n 次压缩视频第 i 个 I 帧第 x 个色度域 CU 的信息标记 $\beta_{n,i,x,\alpha}$，具体表示为

$$\beta_{n,i,x,\alpha}=\begin{cases}1 & \boldsymbol{C}_{n,i,x}\neq\boldsymbol{C}_{n+1,i,x}\text{且 }a_{n,i,x}=\alpha\\ 0 & \text{其他}\end{cases} \tag{7}$$

式中 $\alpha\in\{1,2,\cdots,\varphi\}$。最终根据式(8)计算第 n 次压缩视频所有 I 帧色度域 CU 中选择第 α 种划分类型且属于不同信息 CU 的占比 $W_{n,\alpha}$，具体表示为

$$W_{n,\alpha}=\frac{1}{I_n}\sum_{i=1}^{I_n}\left[\left(\sum_{x=1}^{G_{n,i}}\beta_{n,i,x,\alpha}\right)/H_{n,i,\alpha}\right] \tag{8}$$

式中：I_n、$G_{n,i}$ 分别为第 n 次压缩视频包含的 I 帧数目和第 i 个 I 帧包含的色度域 CU 数目，$H_{n,i,\alpha}$ 为第 n 次压缩视频第 i 个 I 帧中选择第 α 种划分类型的色度域 CU 数目。

下面针对亮度域 CU 的处理与式(3)～(8)所示针对色度域 CU 的处理类似。定义 $\boldsymbol{L}_{n,i,t}$ 为第 n 次压缩视频第 i 个 I 帧第 t 个亮度域 CU 的信息矩阵，具体表示为

$$\boldsymbol{L}_{n,i,t}=[k_{n,i,t},\ m_{n,i,t}] \tag{9}$$

式中 $k_{n,i,t}$、$m_{n,i,t}$ 分别为第 n 次压缩视频第 i 个 I 帧第 t 个亮度域 CU 选择的划分类型及预测模式。它们的取值范围为

$$k_{n,i,t}\in\{1,2,\cdots,\theta\} \tag{10}$$

$$m_{n,i,t}\in\{1,2,\cdots,\omega\} \tag{11}$$

式中 θ 和 ω 分别为 H.266/VVC 标准支持的亮度域 CU 可选的划分类型和预测模式总数。

式(12)为亮度域相同信息 CU 的判别公式，具体表示为

$$\boldsymbol{L}_{n,i,t}=\boldsymbol{L}_{n+1,i,t} \tag{12}$$

对于第 n 次压缩视频第 i 个 I 帧第 t 个亮度域 CU，假如其满足式(12)，则定义其为相同信息 CU，否则为不同信息 CU。然后根据式(13)统计第 n 次压缩视频第 i 个 I 帧第 t 个亮度域 CU 的信息标记 $\delta_{n,i,t,\vartheta}$，具体表示为

$$\delta_{n,i,t,\vartheta}=\begin{cases}1 & \boldsymbol{L}_{n,i,t}\neq\boldsymbol{L}_{n+1,i,t}\text{且 }k_{n,i,t}=\vartheta\\ 0 & \text{其他}\end{cases} \tag{13}$$

式中 $\vartheta\in\{1,2,\cdots,\theta\}$。根据式(14)计算第 n 次压缩视频所有 I 帧亮度域 CU 中选择第 ϑ 种划分类型且属于不同信息 CU 的占比 $V_{n,\vartheta}$，具体表示为

$$V_{n,\vartheta}=\frac{1}{I_n}\sum_{i=1}^{I_n}\left[\left(\sum_{y=1}^{U_{n,i}}\delta_{n,i,t,\vartheta}\right)/S_{n,i,\vartheta}\right] \tag{14}$$

式中 $U_{n,i}$、$S_{n,i,\vartheta}$ 分别为第 n 次压缩视频第 i 个 I 帧包含的亮度域 CU 数目和选择第 ϑ 种划分类型的亮度域 CU 数目[16]。

$W_{n,\alpha}$、$V_{n,\vartheta}$ 为本文构建的高级码流特征。最终基于高级码流特征构建重压缩特征矩阵 $\boldsymbol{F}$，具体表示为

$$\boldsymbol{F}=[W_{n,\alpha},\ V_{n,\vartheta}] \tag{15}$$

3 算法框架

CLF-SVRF 方法框架见图 9，步骤为：

1)对输入的待检测视频码流 B_n，提取所有 I 帧色度域和亮度域 CU 对应的划分类型及预测模式。并分别根据式(3)、(9)构建 B_n 对应的信息矩阵 $\boldsymbol{C}_{n,i,x}$、$\boldsymbol{L}_{n,i,t}$。

2)将 B_n 解码，得到解码视频 Y_n。然后再编码压缩 1 次 Y_n，得到对应的视频码流 B_{n+1}。

3)提取 B_{n+1} 中所有 I 帧色度域和亮度域 CU 对应的划分类型及预测模式。并构建 B_{n+1} 对应的信息矩阵 $\boldsymbol{C}_{n+1,i,x}$、$\boldsymbol{L}_{n+1,i,t}$。

4)根据 $\boldsymbol{C}_{n,i,x}$、$\boldsymbol{L}_{n,i,t}$、$\boldsymbol{C}_{n+1,i,x}$、$\boldsymbol{L}_{n+1,i,t}$，结合式(7)、(8)得到 $W_{n,\alpha}$，结合式(13)、(14)得到 $V_{n,\vartheta}$。并进一步根据式(15)得到 $\boldsymbol{F}$。

5)将 $\boldsymbol{F}$ 输入 SVM 中得到最终的检测结果。

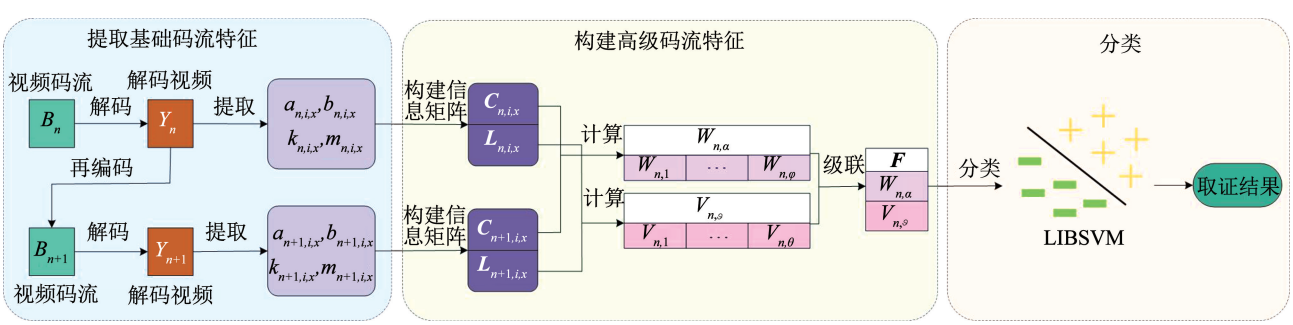


图 9 CLF-SVRF 方法框架

Fig. 9 Framework of CLF-SVRF

在上述步骤中，为了兼顾取证准确率与复杂度，CLF-SVRF 方法只对视频码流编码压缩 1 次，利用第 n 次和第 $n+1$ 次对应的视频码流信息即可完成重压缩取证。另外，第 5 步涉及的 SVM 模型在实际使用前需要提前训练。

4　实验结果

4.1　实验设置

文献[16]提出的方法是当前视频重压缩取证领域代表性的先进方法。本节将 CLF-SVRF 方法与文献[16]方法进行性能比较与分析。选择 68 个包含不同场景内容的 YCbCr 4 :2 :0 格式的监控视频作为测试视频，视频的空间分辨率最小为 720 × 576，最大为 3 840 × 2 160。使用 VTM11.0 编码测试视频，编码帧数为 100 帧，量化参数(quantization parameter，QP)取值如表 3 第 1 列所示，对应于当前视频监控系统普遍使用的码率段。其他编码设置与前节实验的设置一致。首先编码每一个测试视频 1 次生成单压缩视频码流，并标记其为正样本。然后将单压缩视频码流用 VTM11.0 解码得到解码视频，并使用 VTM11.0 将解码视频再编码 1 次得到重压缩视频码流，并将其标记为负样本。

随机选择 41 个测试视频对应的正负样本作为训练集，其他的正负样本作为测试集。训练集用于训练 CLF-SVRF 方法中使用的 SVM 模型。本文采用 LIBSVM 分类器[25-26]，选择径向基函数作为核函数，确定测试集最优取证准确度意义下内核的最佳参数。重压缩取证准确度 A 表示为

$$A=\left(\frac{N_{\kappa,\kappa}}{N_{\kappa,\kappa}+N_{\kappa,\tau}}+\frac{N_{\tau,\tau}}{N_{\tau,\tau}+N_{\tau,\kappa}}\right)/2\times 100\% \quad (16)$$

式中：$N_{\kappa,\kappa}$、$N_{\kappa,\tau}$分别为正样本被判定单压缩视频和重压缩视频的数量；$N_{\tau,\tau}$、$N_{\tau,\kappa}$分别为负样本被判定为重压缩视频和单压缩视频的数量。

4.2　重压缩取证准确度

使用训练集测试文献[16]方法和 CLF-SVRF 方法的重压缩取证准确度，具体结果见表 3。相比于文献[16]方法，CLF-SVRF 方法可以显著提高重压缩取证准确度，对于所有测试视频，重压缩取证准确度平均可以提高 13.53%。为了更深入分析性能，表 4 给出了 8 个代表性视频对应的重压缩取证准确度数据。由表中数据易看出，相比于文献[16]方法，CLF-SVRF 方法对于大部分监控视频具有更高的重压缩取证准确度，但对于个别视频取证准确度未有提升。下面以 City 和 Campus 序列为例分析深层次原因。

表 3　不同方法的重压缩取证准确度

Tab.3　Accuracy of recompression forensic of different methods　%

量化参数	文献[16]方法准确度	CLF-SVRF 方法准确度
12	70.37	81.48
15	66.67	85.19
17	68.52	88.89
19	74.07	92.59
22	70.37	96.30
25	66.67	96.30
27	81.48	94.44
29	85.19	96.30
32	88.89	96.30
35	88.89	100.00
37	92.59	98.15
39	92.59	98.15
42	94.44	92.59
平均	80.06	93.59

表 4　代表性视频的重压缩取证准确度

Tab. 4　Accuracy of recompression forensic of representative videos　%

序列	文献[16]方法准确度	CLF-SVRF 方法准确度
BasketballZoomOut	88.46	96.15
Campus	88.46	84.62
CarparkRotateSlow	84.62	92.31
City	69.23	76.92
Classover	73.08	92.31
DrivingRecorder2	88.46	92.31
videoSRC10	80.77	84.62
WestGateLongView	80.77	100.00

对于 CLF-SVRF 方法，高级码流特征 $W_{n,\alpha}$ 和 $V_{n,\vartheta}$是影响其性能的关键。$W_{n,\alpha}$、$V_{n,\vartheta}$不仅反映了视频码流信息与压缩次数的关系，同时也是 SVM 的输入。图 10 所示量化参数为 25 的情况下 City 和 Campus 序列对应的 $W_{n,\alpha}$和 $V_{n,\vartheta}$。文献[16]方法只提取了亮度域小尺寸 CU 划分类型信息用于构建高级码流特征，如表 2 所示编号分别为 1、6、11 的 4 × 4、8 × 8、16 × 16 的亮度域 CU 划分类型。这种方式对于低分辨率视频通常是适用的，但对于当前普遍使用的大分辨率的内容场景更加丰富的监控视频则不再有效。

对于 City 序列，对比图 10(a)、(b)易得出以下结论：1)相比于 $V_{n,\vartheta}$，$W_{n,\alpha}$对于单压缩视频($n=1$)和重压缩视频($n=2$)的区分度更好，直观表现为图中数据差别更大。这更易使 SVM 训练出性能优异的分割面。2)相比于其他特征，$V_{n,1}$、$V_{n,6}$、$V_{n,11}$对应的重压缩取证区分度较差。这也间接地证明了只提取 4 × 4、8 × 8、16 × 16 的亮度域 CU 划分类型信息很难对 City 序列对应的视频码流进行准确地取证。所以文献[16]方法对于 City 序列的效果较差。对比图 10(c)、(d)易看出，相比于 City 序列，Campus 序列对应的 $W_{n,\alpha}$的重压缩取证区分度较差，且相比于

其他特征，$V_{n,1}$、$V_{n,6}$、$V_{n,11}$ 对应的重压缩取证区分度相对较高。所以对于 Campus 序列，文献[16]方法可以获得更高的重压缩取证准确度。

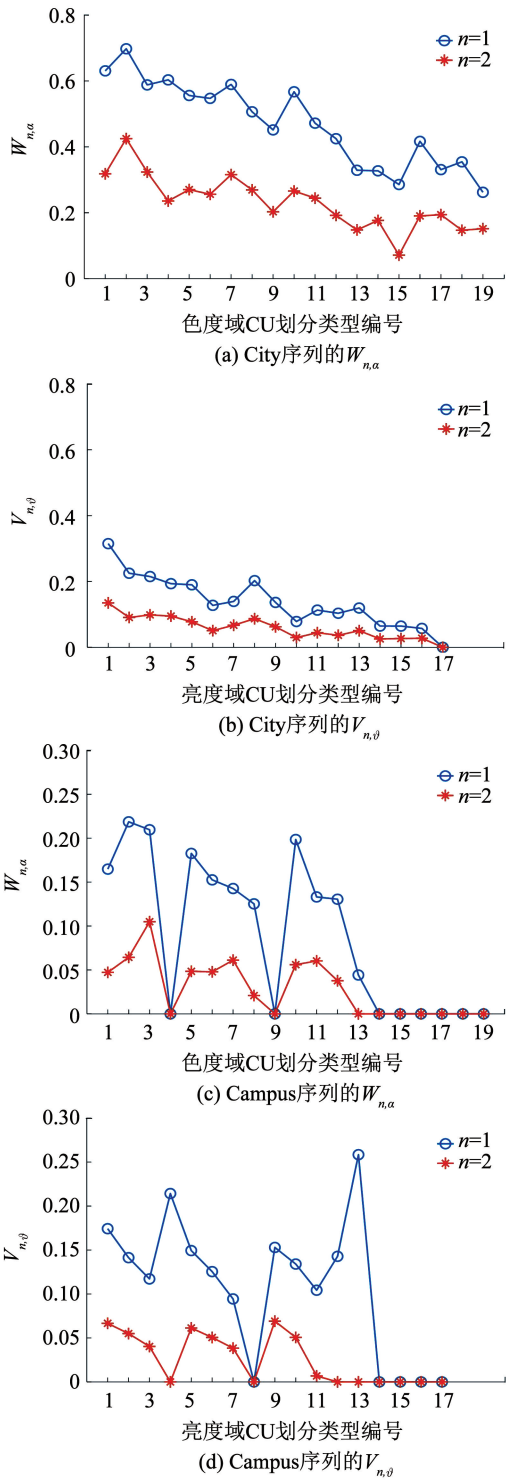


图 10 City 和 Campus 序列的 $W_{n,\alpha}$ 与 $V_{n,\vartheta}$

Fig. 10 $W_{n,\alpha}$ and $V_{n,\vartheta}$ of City and Campus sequences

4.3 复杂度

文献[16]方法和 CLF-SVRF 方法都包括视频编码步骤和其他步骤。一方面，相比于视频解码、构建重压缩特征及分类器检测等步骤，视频编码步骤显著复杂，其占了重压缩取证的大部分时间；另一方面，视频编码是计算机系统独立进行的操作不受人为操作干预。所以本节选择对比文献[16]方法和 CLF-SVRF 方法的编码时间以公正反映其取证复杂度。采用 ψ 衡量相比于文献[16]方法，CLF-SVRF 方法节省编码时间百分比，具体表示为

$$\psi = (|d - \hat{d}|/\hat{d}) \times 100\% \tag{17}$$

式中 d、$\hat{d}$ 分别为 CLF-SVRF 方法、文献[16]方法对应的编码时间。

表 5 为 8 个代表性监控视频的 ψ 数据。由表中数据易看出，相比于文献[16]方法，CLF-SVRF 方法可以显著降低取证复杂度，对于正样本及负样本，编码时间分别可以节省 47.97% 和 46.86%。这主要是因为文献[16]方法需要将待取证的视频码流重新压缩 2 次才能构建出高级码流特征，而 CLF-SVRF 方法只需要将待取证的视频码流重新压缩 1 次即可构建所需的高级码流特征。

表 5 CLF-SVRF 方法节省的编码时间

Tab. 5 Coding time saved by CLF-SVRF %

序列	正样本	负样本
BasketballZoomOut	47.43	47.26
Campus	49.47	46.09
CarparkRotateSlow	45.97	48.23
City	49.38	46.73
Classover	48.96	45.13
DrivingRecorder2	47.80	44.20
videoSRC10	48.14	49.23
WestGateLongView	46.57	48.04
样本平均	47.97	46.86
总平均	47.42	

5 结 论

面向最新的 H.266/VVC 标准，提出了有效融合监控视频色度域和亮度域 CU 划分类型及预测模式信息的视频重压缩取证方法。实验结果证明提出方法具有较高的取证准确度和较低的取证复杂度，更加适用于视频监控应用场景。

由本文研究得出的重要结论包括：对于 H.266/VVC 标准，色度域和亮度域 CU 的划分类型及预测模式信息与视频压缩次数都具有密切关系；监控视频色度域编码信息的应用可以有效提升重压缩取证的准确度；更丰富 CU 划分类型及预测模式的应用可以有效提升重压缩取证方法的普适性；SVM 可以在最新的 H.266/VVC 标准视频重压缩取证中被有效应用。本文只考虑了帧内编码信息，后续考虑帧间编码信息构造基础与高级码流特征是有意义的研究点。

参考文献

[1]任洪昊，朱新山，卢俊彦. 深度图像修复的动态特征融合取证网络[J]. 哈尔滨工业大学学报，2022，54(11)：47
REN Honghao，ZHU Xinshan，LU Junyan. Dynamic feature fusion forensics network for deep image inpainting[J]. Journal of Harbin Institute of Technology，2022，54(11)：47. DOI：10. 11918/202201081

[2]VERDOLIVA L. Media forensics and deepfakes：an overview[J]. IEEE Journal of Selected Topics in Signal Processing，2020，14(5)：910. DOI：10. 1109/JSTSP. 2020. 3002101

[3]WIEGAND T，SULLIVAN G J，BJONTEGAARD G，et al. Overview of the H. 264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology，2003，13(7)：560. DOI：10. 1109/TCSVT. 2003. 815165

[4]SULLIVAN G J，OHM J R，HAN W J，et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology，2012，22(12)：1649. DOI：10. 1109/TCSVT. 2012. 2221191

[5]林敏，林庆毫，翁晓雨，等. 应用于 H. 26X 的通用无损帧内编码优化算法[J]. 哈尔滨工业大学学报，2021，53(8)：7
LIN Min，LIN Qinghao，WENG Xiaoyu，et al. Improved universal lossless intra-frame coding algorithm for H. 26X[J]. Journal of Harbin Institute of Technology，2021，53(8)：7. DOI：10. 11918/202104004

[6]HE Peisong，JIANG Xinghao，SUN Tanfeng，et al. Detection of double compression in MPEG-4 videos based on block artifact measurement[J]. Neurocomputing，2017，228：84. DOI：10. 1016/j. neucom. 2016. 09. 084

[7]BAKAS J，BASHABOINA A K，NASKAR R. MPEG Double compression based intra-frame video forgery detection using CNN[C]//2018 International Conference on Information Technology. Bhubaneswar：IEEE，2018：221. DOI：10. 1109/ICIT. 2018. 00053

[8]ZHANG Zhenzhen，HOU Jianjun，ZHANG Yu，et al. Detecting multiple H. 264/AVC compressions with the same quantisation parameters[J]. IET Information Security，2017，11(3)：152. DOI：10. 1049/iet - ifs. 2015. 0361

[9]FERNÁNDEZ E G，OROZCO A L S，VILLALBA L J G. Digital video manipulation detection technique based on compression algorithms[J]. IEEE Transactions on Intelligent Transportation Systems，2021，23(3)：2596. DOI：10. 1109/TITS. 2021. 3132227

[10]FANG Qianan，JIANG Xinghao，SUN Tanfeng，et al. Detection of HEVC double compression with different quantization parameters based on property of DCT coefficients and TUs[C]//2019 12th International Congress on Image and Signal Processing，BioMedical Engineering and Informatics. Suzhou：IEEE，2019：1

[11]LI Qian，WANG Rangding，XU Dawen. Detection of double compression in HEVC videos based on TU size and quantised DCT coefficients[J]. IET Information Security，2019，13(1)：1. DOI：10. 1049/iet - ifs. 2017. 0555

[12]ELROWAYATI A A，ABDULLAH M F L，ABD MANAF A，et al. Tampering detection of double-compression with the same quantization parameter in HEVC video streams[C]//2017 7th IEEE International Conference on Control System，Computing and Engineering. Penang：IEEE，2017：174

[13]JIA Ruishi，LI Zhaohong，ZHANG Zhenzhen，et al. Double HEVC compression detection with the same QPs based on the PU numbers[C]//ITM Web of Conferences. [S. l.]：EDP Sciences，2016，7：1

[14]XU Qiang，JIANG Xinghao，SUN Tanfeng，et al. Detection of HEVC double compression with non-aligned GOP structures via inter-frame quality degradation analysis[J]. Neurocomputing，2021，452：99. DOI：10. 1016/j. neucom. 2021. 04. 092

[15]UDDIN K，YANG Y，OH B T. Double compression detection in HEVC-coded video with the same coding parameters using picture partitioning information[J]. Signal Processing：Image Communication，2022，103：1. DOI：10. 1016/j. image. 2022. 116638

[16]JIANG Xinghao，XU Qiang，SUN Tanfeng，et al. Detection of HEVC double compression with the same coding parameters based on analysis of intra coding quality degradation process[J]. IEEE Transactions on Information Forensics and Security，2020，15：250. DOI：10. 1109/TIFS. 2019. 2918085

[17]潘鹏飞，姚晔，王慧. 相同编码参数 HEVC 视频重压缩检测[J]. 中国图像图形学报，2020，25(5)：879
PAN Pengfei，YAO Ye，WANG Hui. Detection of double compression for HEVC videos with the same coding parameters[J]. Journal of Image and Graphic，2020，25(5)：879. DOI：10. 11834/jig. 190381

[18]YOUSSEF S，SHANABLEH T. Detecting double and triple compression in HEVC videos using the same bit rate[J]. SN Computer Science，2021，2：406. DOI：10. 1007/s42979 - 021 - 00800 - 8

[19]WU Qiyuan，SUN Tanfeng，JIANG Xinghao，et al. HEVC double compression detection with non-aligned GOP structures based on a fusion feature with optical flow and prediction units[C]//2019 12th International Congress on Image and Signal Processing，BioMedical Engineering and Informatics. Suzhou：IEEE，2019：1

[20]XU Qiang，XU Dongmei，WANG Hao，et al. Detecting double H. 266/VVC compression with the same coding parameters[J]. Neurocomputing，2022，514：231. DOI：10. 1016/j. neucom. 2022. 09. 153

[21]BROSS B，WANG Yekui，YE Yan，et al. Overview of the versatile video coding (VVC) standard and its applications[J]. IEEE Transactions on Circuits and Systems for Video Technology，2021，31(10)：3736. DOI：10. 1109/TCSVT. 2021. 3101953

[22]WU Hongren，REIBMAN A R，LIN Weisi，et al. Perceptual visual signal compression and transmission[J]. Proceedings of the IEEE，2013，101(9)：2025. DOI：10. 1109/JPROC. 2013. 2262911

[23]冯磊，戴明，孙丽娜，等. 利用信息熵判定 HEVC 预测单元尺寸快速算法[J]. 哈尔滨工业大学学报，2016，48(9)：164
FENG Lei，DAI Ming，SUN Lina，et al. An algorithm of fast HEVC PU size decision using information entropy[J]. Journal of Harbin Institute of Technology，2016，48(9)：164. DOI：10. 11918/j. issn. 0367 - 6234. 2016. 09. 028

[24]ITU ISO/IEC. VTM reference software for VVC[EB/OL]. [2020-10-16]. https://vcgit. hhi. fraunhofer. de/jvet/VVCSoftware_VTM/ -/tags

[25]CHANG Chihchung，LIN Chihjen. LIBSVM：A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology，2011，2(3)：1. DOI：10. 1145/1961189. 1961199

[26]HSU Chihwei，CHANG Chihchung，LIN Chihjen. A practical guide to support vector classification[EB/OL]. [2022-10-03]. http://www. csie. ntu. edu. tw/ ~ cjlin/papers/guide/guide. pdf

（编辑　苗秀芝）