

应用于 H.26X 的通用无损帧内编码优化算法

林敏, 林庆毫, 翁晓雨, 陈国捷

(特种光纤与光接入网重点实验室(上海大学), 上海 201900)

摘要: 在 H.26X 系列视频编码标准的无损压缩方案中, 通过帧内预测得到的残差仍具有较强的空间相关性。与自然图像不同, 帧内预测残差的空间相关性体现为含有丰富的边缘特征。为了利用帧内预测残差特殊的空间相关性, 进一步提高视频帧内编码的压缩率, 提出了一种通用的基于残差中值边缘检测的无损帧内编码算法。算法首先对帧内预测残差逐点进行边缘检测, 通过分析临近点的数值特征, 从三种预测模式中选择一种得到当前点的预测值; 然后, 对比预测值与原始残差值, 得到新的预测残差; 最后, 为了确保编码新的预测残差能够提高压缩率, 依照当前编码单元的能量是否降低来快速判断是否使用新的残差进行熵编码。经过该算法处理的编码块具有更低的能量, 从而降低熵编码后的码率。实验结果表明, 在 H.265 与最新的 H.266 标准中应用提出的优化算法, 帧内预测残差的能量平均降低了 67.9%, 平均码率分别降低了 7.04% 和 5.98%, 编解码时间变化细微, 具有显著的实用价值。

关键词: 无损视频压缩; 帧内编码; 预测残差; 高性能视频编码; 多功能视频编码

中图分类号: TN919.81

文献标志码: A

文章编号:

Improved universal lossless intra-frame coding algorithm for H.26X

LIN Min, LIN Qinghao, WENG Xiaoyu, CHEN Guojie

(Key Laboratory of Specialty Fiber Optics and Optical Access Network (Shanghai University), Shanghai 201900, China)

Abstract: In lossless compression scheme of the H.26X series video coding standards, residuals obtained by intra-frame prediction still retain strong spatial correlation. Unlike natural images, intra-frame prediction residuals contain rich edge features. To take advantage of residuals' special spatial correlation, further improve the intra-frame coding compression ratio, a universal lossless intra-frame coding algorithm based on residual median edge detection is proposed in this paper. Algorithm first performs edge detection on the prediction residuals sample-by-sample, by analyzing the numerical features of the adjacent residuals, select a prediction mode to obtain new prediction value in three modes; then, new prediction residual is obtained by comparing prediction value and original residual; Finally, according to whether the energy of the current coding unit is reduced, it can quickly determine whether to use new residuals for entropy coding. Coding blocks processed by this algorithm have lower energy, thereby reducing the bit-rate after entropy coding. Experimental results show that applying the proposed improved algorithm in H.265 and the latest H.266 standards, the average energy of residuals decrease by 67.9%, the average bit-rate is reduced by 7.24% and 6.16%, respectively, while the codec time changes slightly, with significant practical value.

Key Words: lossless video compression; intra-frame coding; prediction residuals; HEVC; VVC

高级视频编码 H.264/AVC (Advanced Video Coding)、高性能视频编码 H.265/HEVC (High Efficiency Video Coding)、多功能视频编码 H.266/VVC (Versatile Video Coding) 是由国际电信联盟与国际标准化组织联合专家组制定的一系列视频压缩编码标准^[1-3], 代表着自 2003 年至今各时期最先进的视频编码技术。帧内编码是 H.26X 系列标准中的重要组成部分, 其利用图像的空间相关性进行数据压缩, 视频的关键帧也只使用帧内编码处理。加之近年来, 自动驾驶视觉、云游戏、远程桌面共享、后期制作、医学影像存档等应用的发展使得视频无损帧内编码的需求日益增多^[4]。因而探索无损帧内编码的优化算法有很高的研究和应

收稿日期: 2021-mm-dd

基金项目: 国家重点研发计划(2019YFB2204500)

作者简介: 林敏 (1974—), 男, 教授, 博士生导师

通信作者: 林敏, mlin99@shu.edu.cn

用价值。

H.26X 系列标准可以通过简单地跳过变换、量化、去块滤波、自适应样点补偿等可能引入失真的步骤，实现无损帧内压缩^[5]。但也因此使得待编码系数具有较高的能量，为后续的熵编码带来了极大的压力。为了优化 H.26X 系列标准无损帧内编码的性能，国内外学者进行了大量的研究。Kamisli^[6]提出了一种不会引入失真的整数离散正弦变换，一定程度上弥补了跳过变换带来的码率损失；Sanchez 等^[7]针对不同的残差数值分布特征设计了 3 组映射规则，整体降低了残差的能量；残差差分脉冲编码调制（Residual Differential Pulse Code Modulation, RDPCM）是 HEVC 屏幕图像编码扩展标准（HEVC Screen Content Coding Extension, HEVC-SCC）^[8]的一部分，通过对残差进行再处理的方式提高了编码效率；区块差分脉冲编码调制（Block Differential Pulse Code Modulation, BDPCM）^[1]是 H.266 标准的一部分，通过在水平或垂直方向上使用临近像素迭代预测的方式提高了预测准确性；Zhou 等^[9,10]将 DPCM 的思想扩展到了任意预测角度，提出了基于样本点的帧内角度预测（Sample-based Angular Intra-Prediction, SAP），进一步提高了预测准确性；元辉等^[11]提出了一种新颖的不需要编码预测模式信息的帧内预测方案；Zhang 等^[12]分析了亮度通道与色差通道之间的关系，设计了 3 个模型，使用亮度通道对色差通道进行帧内预测；Li 等^[13]通过将参考像素拓展到多行的方式提高了帧内预测的准确性。

上述文献中提出的算法对帧内编码的性能均有不同程度的优化，但仍存在以下待改进的问题：[6]设计的整数变换方法局限性很大，且没有量化的配合很难降低待编码系数的能量，因此带来的码率优化较小；[7,8]设计的残差分段映射方案没有考虑残差独特的空间相关性，仍有优化空间；[1,9-13]都是设法增加或优化预测模式来提高帧内预测的准确性，然而各代 H.26X 标准中的帧内预测模式不断增多，在最新的 H.266 中已经达到 67 种，很难再挖掘出优化空间，可以预见今后这类改进带来的优化会越来越小；最后，上述大部分方案具有一定的应用范围的局限性，很难在各世代的 H.26X 标准中通用。基于上述原因，本文提出了一种在 H.264、H.265、H.266 中通用的无损帧内编码优化算法。该算法利用了帧内预测残差具有丰富的边缘特征这一统计结果，对预测残差进行处理。首先将当前编码块的残差首行、首列作为参考点；然后利用中值边缘检测（Median Edge Detection, MED）对剩余点逐点进行预测并求得新的残差；最后统计处理前后的残差能量，选择能量较小的一组进行熵编码。算法经过测试，得到了明显的码率优化，且该算法具有极强的通用性，可在各世代的 H.26X 标准中应用，或与其他优化方法结合应用，进一步提高优化效果。

1 H.26X 无损帧内编码

H.26X 系列标准中的帧内编码方案经过了各世代的持续优化，但整体框架仍保持一致，包含帧内预测-变换-量化-熵编码 4 个基本流程（图 1 中 a）。帧内编码过程中引入失真的是变换和量化操作，因此可通过跳过变换和量化，仅保留帧内预测和熵编码来实现基础的 H.26X 无损帧内编码^[5]（图 1 中 b）。

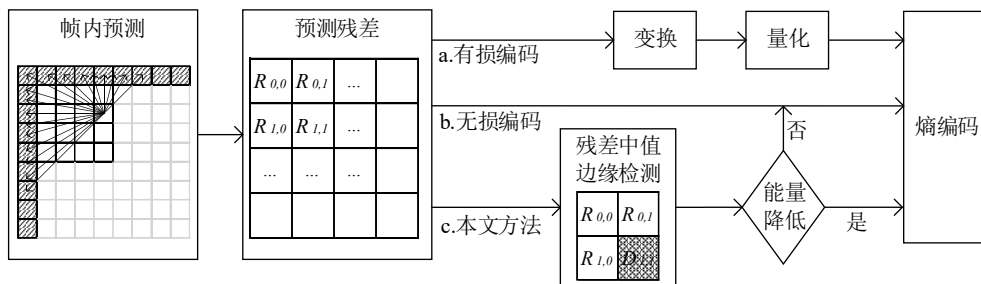


图 1 H.26X 帧内编码流程

Fig.1 H.26X intra-frame coding process

帧内预测是 H.26X 无损帧内编码的核心部分。对每个预测单元（Prediction Unit, PU），H.26X 执行帧内预测得到相应的预测残差 $R_{i,j}$ 为

$$R_{i,j} = P_{i,j} - I_{i,j}, \quad (1)$$

式中 $I_{i,j}$ 表示原始像素值， $P_{i,j}$ 表示通过参考像素结合不同的帧内预测方式得到的预测值。H.26X 的帧内预测模式包括平滑预测（DC 模式、Planar 模式）和方向预测两大类，随着编码标准的发展，预测模式的数量不断增多。H.264 仅使用 9 种预测模式^[3]，H.265 增加到 35 种^[2]，H.266 达到了 67 种^[1]。

图 2 展示了在 H.265 标准下经过帧内预测后的残差图像（数值整体平移了 128 以绘制负值）。出于多种原因，尽管帧内预测的准确性不断提高，残差中仍然可能保留丰富的边缘特征。首先，由于 PU 最小尺寸的限制，角度预测在图像中纹理丰富的区域始终无法得到良好的预测结果；其次，当被预测块靠近参考像素的边沿存在不连续性时，角度预测很可能会插入原始块中本不存在的方向性条纹。

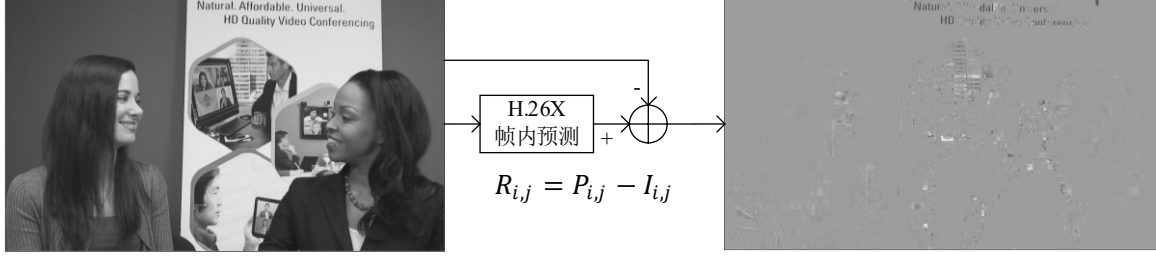


图 2 “KristenAndSara” 第一帧及其帧内预测的残差（亮度）

Fig.2 First frame of “KristenAndSara” and its intra-frame prediction residual (luma)

2 基于残差中值边缘检测的无损帧内编码

针对预测残差图像中存在丰富的边缘特征这一统计结果，本文提出了基于残差中值边缘检测（Residual Median Edge Detection, R-MED）的无损帧内编码算法。算法通过对预测残差进行二次预测，有效地压缩了残差图像中丰富的边缘特征，从而降低待编码系数的整体能量，最终降低视频码率。

2.1 残差中值边缘检测算法

MED 已被应用在图像的低复杂度无损压缩（Low Complexity Lossless Compression for Images, LOCO-I）^[7,14]算法中。但与 LOCO-I 使用 MED 对原始像素进行预测不同，本文考虑到 MED 具有很强的边缘检测能力，恰好残差图像中包含丰富的边缘特征，因此将其应用到残差图像的二次预测上。R-MED 的新预测值 $P'_{i,j}$ 由下式给出：

$$P'_{i,j} = \begin{cases} \min(R_{i,j-1}, R_{i-1,j}), & R_{i-1,j-1} > \max(R_{i,j-1}, R_{i-1,j}); \\ \max(R_{i,j-1}, R_{i-1,j}), & R_{i-1,j-1} < \min(R_{i,j-1}, R_{i-1,j}); \\ R_{i,j-1} + R_{i-1,j} - R_{i-1,j-1}, & \text{Otherwise.} \end{cases} \quad (2)$$

式中 $R_{i,j}$ 表示 H.26X 经过帧内预测后得到的残差。如图 3 所示，R-MED 可以精准地得到强边缘处的预测值。当 $R_{i-1,j-1} > \max(R_{i,j-1}, R_{i-1,j})$ 时（见图 3（a）），如果 $R_{i,j-1} > R_{i-1,j}$ ，那么在 $P'_{i,j}$ 左侧的垂直方向上很可能存在明显的垂直边缘，此时最佳预测值为 $R_{i-1,j}$ ；反之则很可能在上方存在明显的水平边缘，此时最佳预测值为 $R_{i,j-1}$ 。类似地，图 3（b）展示了 $R_{i-1,j-1} < \min(R_{i,j-1}, R_{i-1,j})$ 时， $P'_{i,j}$ 的选择依据。

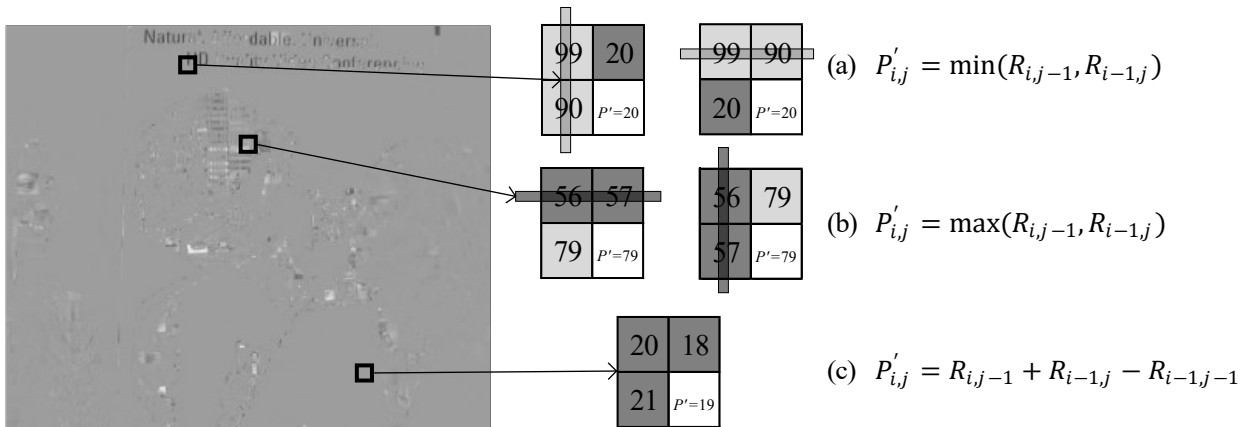


图 3 R-MED 的预测策略

Fig.3 Prediction strategy of R-MED

此外，如果 R-MED 没有检测到明显的水平或垂直边缘，将采取模仿水平/垂直方向相邻点变化趋势的预测策略。如图 3（c）所示， $P'_{i,j}$ 周围的参考点并不满足式（2）的前两个条件，表明该区域变化平坦没有

复杂纹理。 $P'_{i,j}$ 上方参考点的数值由 20 降低到 18 (-2)，左侧参考点的数值由 20 增加到 21 (+1)，因此 $P'_{i,j}$ 模仿上方的变化趋势得到 $P'_{i,j} = R_{i,j-1} - 2 = 21 - 2 = 19$ ，或模仿左侧的变化趋势得到 $P'_{i,j} = R_{i-1,j} + 1 = 18 + 1 = 19$ 。

2.2 R-MED 算法应用

通过上述分析，证明了 R-MED 具有对边缘信息预测准确的特点，因此适合将其应用到 H.26X 残差图像的二次预测中。如图 1 中 c 所示，本文对经过 H.26X 帧内预测后，进行熵编码前的预测残差进行再处理。以 PU 为基本单位，将各 PU 的首行、首列作为参考点，对剩下的数据点逐点应用 R-MED 算法并计算新的残差 $D_{i,j}$ ：

$$D_{i,j} = \begin{cases} R_{i,j}, & i = 0 \text{ or } j = 0; \\ P'_{i,j} - R_{i,j}, & \text{Otherwise.} \end{cases} \quad (3)$$

式中 $R_{i,j}$ 表示 H.26X 经过帧内预测后得到的残差， $P'_{i,j}$ 表示通过 R-MED 算法得到的预测值。然后分别计算处理前后的残差总能量 E_{src} 、 E_{R-MED} ：

$$E_{src} = \sum_i \sum_j R_{i,j}^2, \\ E_{R-MED} = \sum_i \sum_j D_{i,j}^2,$$

最后选择总能量较小的一组进行熵编码。通过残差的总能量判断是否选用经过二次处理的残差是因为，H.26X 在对预测残差进行编码前，需要使用哥伦布-莱斯码或指数哥伦布码对其进行二值化，二值化后码字的长度将与残差的绝对幅值呈指数关系^[15]。图 4 是一个 4×4 的预测单元经过 H.265 + R-MED 算法处理前后的效果。

0	0	-2	-1
-1	-1	-2	-1
0	1	0	0
0	-1	-2	-1

$E_{src} = 19$

0	0	-2	-1
-1	0	0	0
0	-1	0	0
0	2	1	-1

$E_{R-MED} = 13$

图 4 一个经过 R-MED 处理的预测单元

Fig.4 A prediction unit processed by R-MED

为了验证 R-MED 算法的有效性，统计了在 H.265 标准下，部分测试序列第 1 帧经过算法处理后整体能量的变化率。统计结果如表 1 所示，同时绘制测试序列“KristenAndSara”第一帧经过算法处理前后的残差图像（数值整体平移了 128 以绘制负值），如图 5 所示。统计结果显示单帧图像的残差能量平均降低了 67.9%，观察“KristenAndSara”第一帧的残差图像发现边缘特征极大减少，证明 R-MED 算法对预测残差图像有很准确的预测能力。

表 1 应用 H.265 + R-MED 算法后残差能量的变化率

Tab.1 Reduced ratio of residual energy processed by H.265+R-MED

类别	序列	分辨率	原始残差总能量	R-MED 处理后残差总能量	变化率/%
A	PeopleOnStreet	2560×1600	633451646	94108069	-85.1
B	Kimono	1920×1080	73752438	11522730	-84.4
C	BQMall	832×480	59358254	25243957	-57.5
D	BlowingBubbles	416×240	14714641	8036874	-45.4
E	KristenAndSara	1280×720	97320940	12361610	-87.3
F	SlideEditing	1280×720	1133005976	595102758	-47.5
AVG	-	-	-	-	-67.9

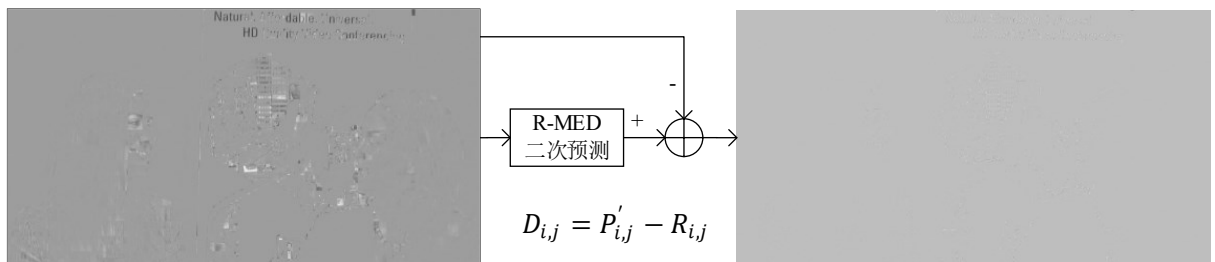


图 5 “KristenAndSara”的预测残差经过 R-MED 处理后的结果（亮度）

Fig.5 Output of prediction residual of “KristenAndSara” after R-MED processing (luma)

3 实验结果与分析

R-MED 算法分别在 H.265 参考软件 HM-SCC-16^[16]和 H.266 参考软件 VTM-12^[17]中实现,并在视频联合专家组制定的公共测试条件下使用全帧内 (All Intra, AI) 配置进行实验。为了证明算法的适用性,测试序列使用了专家组推荐的 ClassA-ClassF 共 6 类 25 个序列,测试序列的相关属性及 R-MED 算法对比 H.265 及 H.266 的测试结果如表 2 所示:

表 2 R-MED 算法应用在 H.265 及 H.266 的性能测试

Tab.2 Performance test of R-MED algorithm applied to H.265 and H.266

序列类别	序列	分辨率	帧数	H.265 + R-MED vs		H.266 + R-MED vs	
				H.265 bypass	H.265	H.266 bypass	H.266
				TransQuant/%	RDPCM ^[8] /%	TransQuant/%	BDPCM ^[1] /%
A	PeopleOnStreet	2560×1600	150	-11.10	-5.99	-9.67	-5.99
A	Traffic	2560×1600	150	-10.10	-3.54	-8.60	-3.77
B	BasketballDrive	1920×1080	500	-2.19	-0.49	-1.42	-0.52
B	BQTerrace	1920×1080	600	-5.43	-0.96	-4.16	-1.15
B	Cactus	1920×1080	500	-2.92	-0.99	-2.14	-1.05
B	Kimono	1920×1080	240	-6.82	-2.63	-5.96	-3.35
C	ParkScene	832×480	240	-5.85	-2.26	-4.87	-2.58
C	BasketballDrill	832×480	500	-2.61	-1.14	-1.92	-1.22
C	BQMall	832×480	600	-4.16	-0.66	-3.10	-0.86
C	PartyScene	832×480	500	-3.80	-0.51	-2.65	-0.69
D	RaceHorsesC	416×240	300	-7.03	-2.64	-6.13	-3.25
D	BasketballPass	416×240	500	-9.10	-1.92	-7.71	-2.69
D	BlowingBubbles	416×240	500	-4.11	-1.93	-3.15	-0.75
D	BQSquare	416×240	600	-3.10	-0.48	-2.24	-0.56
D	RaceHorses	416×240	300	-8.29	-3.53	-7.23	-3.58
E	FourPeople	1280×720	600	-9.23	-3.83	-7.92	-3.97
E	Johnny	1280×720	600	-6.99	-2.58	-6.13	-3.04
E	KristenAndSara	1280×720	600	-7.35	-2.60	-6.40	-2.82
E	Vidyo1	1280×720	600	-9.93	-3.30	-8.56	-3.65
E	Vidyo3	1280×720	600	-9.10	-3.60	-7.64	-3.80
E	Vidyo4	1280×720	600	-8.92	-2.65	-7.36	-3.04
F	BasketballDrillText	832×480	500	-3.01	-1.19	-2.21	-1.17
F	ChinaSpeed	1024×768	500	-9.13	-1.35	-7.86	-1.47
F	SlideEditing	1280×720	300	-7.33	-0.51	-7.20	-0.80
F	SlideShow	1280×720	500	-18.45	-5.16	-17.36	-6.73
平均码率优化		-	-	-7.04	-2.26	-5.98	-2.50
编码时间		-	-	104	107	117	111
解码时间		-	-	95	98	83	93

与 H.265 参考软件 HM-16 相比,经过 R-MED 算法处理的视频序列达到了最大 18.45%,平均 7.04%的码率优化,平均编码时间仅增加 4%,同时由于经过处理的待编码残差能量大幅减少,缓解了后续熵编解码器的压力,因此平均解码速度加快 5%,另外还证明了比 HEVC-SCC 扩展标准中的 RDPCM 性能更好;与 H.266 参考软件 VTM-12 相比,视频码率达到了最大 17.36%,平均 5.98%的优化,编码时间平均增加 17%,解码速度提升 17%,同时证明了比 H.266 标准中的 BDPCM 性能更好。

4 结语

针对 H.26X 系列编码标准中,无损帧内预测得到的残差仍保留特殊的空间相关性,整体待编码残差能量较高的问题,本文提出了基于残差中值边缘检测的无损帧内编码优化算法。算法通过对预测残差进行边缘检测得到高准确度的预测值,由此计算得到新的能量大幅降低的残差值。在 H.265 与最新的 H.266 标准的参考软件中应用提出的优化算法,平均码率分别降低了 7.04%和 5.98%,编解码时间变化细微,具有显著的实用价值。同时该算法具有很强的通用性,可在 H.26X 系列标准中应用,也可与各种帧内预测优化方法结合使用,甚至有可能在编码结构相似的数字音视频编码技术标准(Audio Video Standard, AVS)、AV1(AOMedia Video 1)标准中应用。不足的是,文中描述的算法暂时只能应用在无损编解码中,今后的工作将研究 R-MED 算法与其他优化算法的联合应用,并将该方法拓展到无损编码、帧间编码。

参考文献:

- [1] WIEN M, BROSS B. Versatile Video Coding – Algorithms and Specification[C]//2020 IEEE International Conference on Visual Communications and Image Processing (VCIP). Macau: IEEE, 2020: 1. DOI: 10.1109/VCIP49819.2020.9301820
- [2] 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
- [3] MARPE D, WIEGAND T, SULLIVAN G J. The H.264/MPEG4 advanced video coding standard and its applications[J]. IEEE Communications Magazine, 2006, 44(8): 134. DOI: 10.1109/MCOM.2006.1678121
- [4] DE-LUXÁN-HERNÁNDEZ S, VENUGOPAL G, GEORGE V, et al. A Fast Lossless Implementation Of The Intra Subpartition Mode For VVC[C]//2020 IEEE International Conference on Image Processing (ICIP). Abu Dhabi: IEEE, 2020: 1118. DOI: 10.1109/ICIP40778.2020.9191103
- [5] MRAK M, XU J. Improving screen content coding in HEVC by transform skipping[C]//2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO). Bucharest: IEEE, 2012: 1209
- [6] KAMISLI F. Lossless Image and Intra-Frame Compression With Integer-to-Integer DST[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(2): 502. DOI: 10.1109/TCSVT.2017.2787638
- [7] SANCHEZ V, AULÍ-LLINÀS F, SERRA-SAGRISTÀ J. Piecewise Mapping in HEVC Lossless Intra-Prediction Coding[J]. IEEE Transactions on Image Processing, 2016, 25(9): 4004. DOI: 10.1109/TIP.2016.2571065
- [8] XU J, JOSHI R, COHEN R A. Overview of the Emerging HEVC Screen Content Coding Extension[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 50. DOI: 10.1109/TCSVT.2015.2478706
- [9] ZHOU M, GAO W, JIANG M, et al. HEVC Lossless Coding and Improvements[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1839. DOI: 10.1109/TCSVT.2012.2221524
- [10] SANCHEZ V, LLINÀS F A, RAPESTA J B, et al. Improvements to HEVC Intra Coding for Lossless Medical Image Compression[C]//2014 Data Compression Conference. Snowbird: IEEE, 2014: 423. DOI: 10.1109/DCC.2014.76
- [11] 元辉, 常义林, 卢朝阳, 等. 一种降低预测模式开销的帧内预测方法[J]. 西安电子科技大学学报, 2010, 37(6): 981. DOI: 10.3969/j.issn.1001-2400.2010.06.001
YUAN Hui, CHANG Yilin, LU Zhaoyang, et al. Intra prediction method for reducing prediction mode information[J]. Journal of Xidian University, 2010, 37(6): 981. DOI: 10.3969/j.issn.1001-2400.2010.06.001
- [12] ZHANG K, CHEN J, ZHANG L, et al. Enhanced cross-component linear model for chroma intra-prediction in video coding[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3983. DOI: 10.1109/TIP.2018.2830640
- [13] LI J, LI B, XU J, et al. Efficient Multiple-Line-Based Intra Prediction for HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(4): 947. DOI: 10.1109/TCSVT.2016.2633377
- [14] WEINBERGER M J, SEROUSSI G, SAPIRO G. The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS[J]. IEEE Transactions on Image Processing, 2000, 9(8): 1309. DOI: 10.1109/83.855427
- [15] SOLE J, JOSHI R, NGUYEN N, et al. Transform Coefficient Coding in HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1765. DOI: 10.1109/TCSVT.2012.2223055
- [16] BOSSEN F, SUEHRING K, IWAMURA S, et al. HEVC Reference Software HM-16[CP/OL]. [2021-03-18].

<https://vcgit.hhi.fraunhofer.de/jvet/HM>

- [17] BOSSEN F, GALLASSO M P, WIECKOWSKI A, et al. VVC Reference Software VTM-12.0[CP/OL]. [2021-02-01].
https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM