



Complexity assessment of the intra prediction in Versatile Video Coding

Naima Zouidi, Amina Kessentini, Wassim Hamidouche, Nouri Masmoudi,
Daniel Menard

► To cite this version:

Naima Zouidi, Amina Kessentini, Wassim Hamidouche, Nouri Masmoudi, Daniel Menard. Complexity assessment of the intra prediction in Versatile Video Coding. Multimedia Tools and Applications, 2023, 82, pp.27751-27770. 10.1007/s11042-023-14442-w . hal-04056490

HAL Id: hal-04056490

<https://hal.science/hal-04056490v1>

Submitted on 12 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Complexity assessment of the intra prediction in Versatile Video Coding

Naima Zouidi* · Amina Kessentini · Wassim Hamidouche · Nouri Masmoudi · Daniel Menard

Received: date / Accepted: date

Abstract The Versatile Video Coding (VVC) was finalized in July 2020 by the Joint Video Experts Team (JVET). During the standardization process, many coding techniques were investigated in order to provide superior coding efficiency with high level of versatility over the High Efficiency Video Coding (HEVC). Indeed, VVC achieved about 40% of bitrate saving compared to its predecessor for the same objective visual quality. However, its encoding complexity has considerably increased due to the exhaustive Rate-Distortion (RD) search. In this paper, a complexity assessment of the intra mode decision of VVC is presented. Experimental results reveal that the intra coding tools and the steps of intra mode decision have various degrees of complexity. Thus, opening up several complexity reduc-

tion opportunities, that yield up to 43% of complexity reduction. Also, VVC tends usually to select the 6 Most Probable Modes (MPMs). However, this selection may depend on the used intra coding tool and video texture characteristics. These conclusions give clear openings to accelerate the intra mode decision.

Keywords Versatile Video Coding · Complexity reduction · Intra prediction · Rate Distortion · Intra mode decision

1 Introduction

Digital video is ushering in a new era of evolution. In fact, new video formats such as High Dynamic Range (HDR), 360 ° /Virtual Reality (VR), and Augmented Reality (AR) have emerged, which should allow real progress in video streaming and broadcasting. Nevertheless, these video formats are extremely heavy to deliver, as they require a huge amount of raw data. At once, video traffic is continuing to grow at a very rapid pace. It should represent around 82% of global internet traffic by 2022 [1] and thus deepening the bandwidth limitation issue. Since the late 1980s, digital video coding standards were recognized to provide efficient video storage and transmission [2]. As an example, the High Efficiency Video Coding (HEVC), established in 2013, has doubled the coding efficiency beyond prior standards with the aim of enabling full High Definition (HD) video transmission. However, as video resolution will increase from HD to Ultra-High Definition (UHD) and new highly intensive data applications are appearing, this coding efficiency is no longer sufficient. Consequently, the Joint Video Experts Team (JVET), created jointly between the International Telecommunications Union (ITU) and the International Organization for

N. Zouidi
ENI Sfax - LETI and INSA Rennes - IETR,
20 Avenue of Buttes de Coesmes, Rennes cedex 7, France,
35708
E-mail: Naima.zouidi@insa-rennes.fr

A. Kessentini
ENI Sfax - LETI and ISIM Gabes,
Road of Soukra, Sfax, Tunisia, 3038

W. Hamidouche
INSA Rennes - IETR,
20 Avenue of Buttes de Coesmes, Rennes cedex 7, France,
35708

N. Masmoudi
ENI Sfax - LETI,
Road of Soukra, Sfax, Tunisia, 3038

D. Menard
INSA Rennes - IETR,
20 Avenue of Buttes de Coesmes, Rennes cedex 7, France,
35708

Standardization (ISO), released in 2020, the Versatile Video Coding (VVC). Since coding efficiency presents a critical enabler of video sharing, the standardization bodies have focused on halving the bitrate requirement of the VVC compared to HEVC with the same objective visual quality. As of the VVC Test Model (VTM)-8.0, bitrate savings of 24% in the All Intra (AI) and 35% in the Random Access (RA), respectively [3] were achieved beyond its predecessor HEVC Test Model (HM).

One of the main contributors to this achievement is the new block partitioning structure namely, Quadtree with Nested Multi-type Tree (QTMT) [4]. This new structure enables better adaptation to the local activity of the pixels. Furthermore, the intra prediction have been improved by extending existing tools or introducing new ones. In fact, the number of angular modes has been increased and new tools such as Intra Sub-Partitions (ISP) [5], Multi-Reference Lines (MRL) [6], and Matrix weighted Intra Prediction (MIP) [7] have been introduced. Yet, regarding complexity, the VTM encoder is $31\times$ more complex than the HM encoder in the AI configuration [8].

This complexity could primarily affect the real-time implementation of the VVC codec. In other words, the new codec deployment would require intensive hardware support, that most devices aren't able to provide. Regarding this issue, some works have considered complexity assessments and in-depth analyses of the VVC in order to quantify its complexity and guide the video coding community in developing new complexity reduction techniques. In [9] and [10], complexity assessment of the VVC codec compared to HEVC was conducted under several coding conditions. Hence, in [9], the effect of numerous encoding aspects on the VVC encoder runtime was studied. Moreover, in [10], the complexity of the VVC encoder and decoder compared to the HEVC was highlighted through a profiling of encoder/decoder runtimes and memory usage. Previous work [11], characterized the upper bound of complexity reduction in the VVC encoder. It considered different combinations of the three key levels of the exhaustive Rate-Distortion (RD) search: block partitioning, intra prediction and transform selection. In addition, other works [12], [13] and [14] proposed fast intra mode decision algorithms based on in-depth analyses of this coding step. For instance, authors in [12] and [14] studied the correlation between the texture information of Coding Unit (CU) and the best intra mode. Whereas, in [13] authors identified the relationship between the best mode in Rough Mode Decision (RMD) and the Most Probable Modes (MPMs). Also, previous works in [15] and [16], studied the intra mode decision according to several encoding parameters and makes full use of several relations to

speed up the VVC encoder. In [17], the advantages of using Single Instruction on Multiple Data (SIMD) optimization for VVC was studied and the new coding tools were profiled to identify the most time-demanding ones. Furthermore, [18] introduces a comparative analysis between VVC and HEVC codecs in order to evaluate the VVC coding efficiency and complexity overhead. In [19] a rate-distortion-complexity study of the new intra coding tools over different quantization scenarios was conducted. Finally, authors in [20] reviewed the new intra coding tools of VVC and gave a subjective and an objective studies of these latter with respect to the HEVC.

In this paper, a complexity assessment of the VVC intra prediction is introduced. While previous works analyzed the complexity of the VVC encoder and the upper bound of complexity reduction in general, our paper gives a deeper complexity assessment of the intra prediction in VVC. It studies the complexity of the new intra coding tools and their impact on coding efficiency. Then, the complexity of the intra mode decision steps is highlighted through a profiling assessment and the upper bound of complexity reduction in each of these decision steps is emphasized, including how they effect the intra mode decision runtime when they are disabled. Finally, this paper explores the most time-demanding step of the intra mode decision. Hence, studying the Final Intra Prediction Mode Decision (FIPMD) according to used intra coding tool and video texture characteristics.

The remainder of this paper is organized as follows. Section 2 reviews the key tools of the VVC, focusing on the new intra coding tools and the intra mode decision steps. Then, a complexity assessment of the VVC intra prediction is conducted in Section 3. Section 4 analyzes the upper bound of complexity reduction in each of the intra mode decision steps and their impact on this decision runtime when they are disabled. Next, Section 5 provides an in-depth statistical analysis of the selected Intra Prediction Modes (IPMs). Finally, Section 6 concludes this paper.

2 Brief overview of the VVC standard

This section gives a brief overview of the key tools of the VVC standard with a particular focus on its intra coding tools and intra mode decision steps.

2.1 Key tools of the VVC

In response to the issued call for proposal, there are several contributions to the Versatile Video Coding (VVC)

that have been investigated. For example, the Quadtree with Nested Multi-type Tree (QTMT) [4] was introduced for block partitioning. The QTMT concept consists in nesting a Multi-type Tree (MT) to the Quadtree (QT) of the High Efficiency Video Coding (HEVC) in order to make the coding process more flexible to forthcoming video resolutions and contents. Right after applying the QT partitioning to the Coding Tree Unit (CTU), the MT further partition the QT leaf nodes into 2 or 3 partitions vertically or horizontally [4]. This partitioning process is based on the partitioning mode, either binary or ternary. Fig 1 illustrates the different partitioning modes in the QTMT.

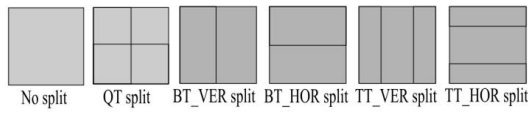


Fig. 1: QTMT partitioning modes

Despite using multiple processing units during partitioning, prediction, and transformation, only one processing unit is used in VVC for the entire coding process. This latter is referred to as Coding Unit (CU) and it may have either square or rectangular shape [21]. Since, the directional intra prediction plays the main role in efficiently predicting the different directional structures in video content, the number of angular Intra Prediction Modes (IPMs) was extended from 33 to 65. This new set of IPMs still covers the same angular range as that of HEVC. It starts from 45° to -135° , where modes 2 and 66 indicate the angle 45° and -135° , respectively. In Fig. 2, the 35 original IPMs of HEVC are illustrated as black arrows. Whereas, the new IPMs in VVC are illustrated as green dotted arrows.

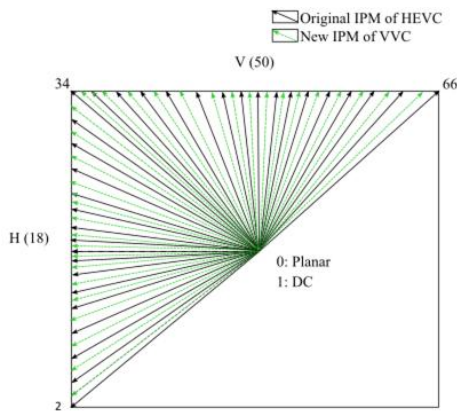


Fig. 2: Illustration of the 67 IPMs in VVC. V and H are the vertical and horizontal modes, respectively

In order to increase the coding efficiency of intra prediction, several intra coding tools were also introduced. For the 67 regular modes (DC, planar, and the 65 angular modes with no intra coding tools), the intra prediction is carried-out from neighbouring pixels, that correspond to reference line 0 as depicted in Fig. 3. The Multi-Reference Lines (MRL) tool allows using farther reference lines [5], as shown in Fig. 3, to predict the current CU. In this case, reference lines 1 or 3 can be considered. The index of reference lines is then signaled with the MRL flag.

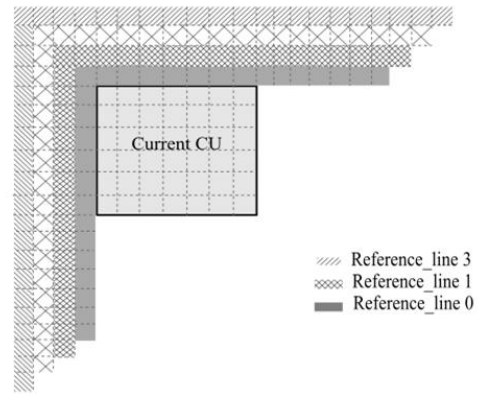


Fig. 3: Reference lines neighbouring to an intra-coded block

The MRL tool is applied only on the 6 Most Probable Modes (MPMs), which are created based on the modes of neighbouring blocks [21], as detailed in Fig. 4. The modes m_L and m_A are the modes of left and above neighbouring blocks, respectively. Hence, when m_L and m_A are non-angular, then DC, planar with horizontal (H), vertical (V) modes and their derived modes are considered during the creation of the 6 MPMs. However, if these latter are both angular and the same, then only derived modes from m_L are considered. Otherwise, the maximum mode m_{max} and the minimum mode m_{min} between m_L and m_A , can be considered with their derived modes too.

In order to adapt the intra prediction to various texture characteristics, the VVC introduced the Intra Sub-Partitions (ISP) tool. This tool enables the partitioning of intra-coded block vertically or horizontally into 2 or 4 sub-partitions, in which the same IPM is used. Hence, for each mode, there are two options, when ISP is enabled, either partitioning the intra-coded block in a horizontal way (ISP = 1) or in a vertical way (ISP = 2). Furthermore, each sub-partition should be predicted using the reconstructed samples of the previously coded sub-partition [6].

Besides the simple weighted intra prediction [22], a

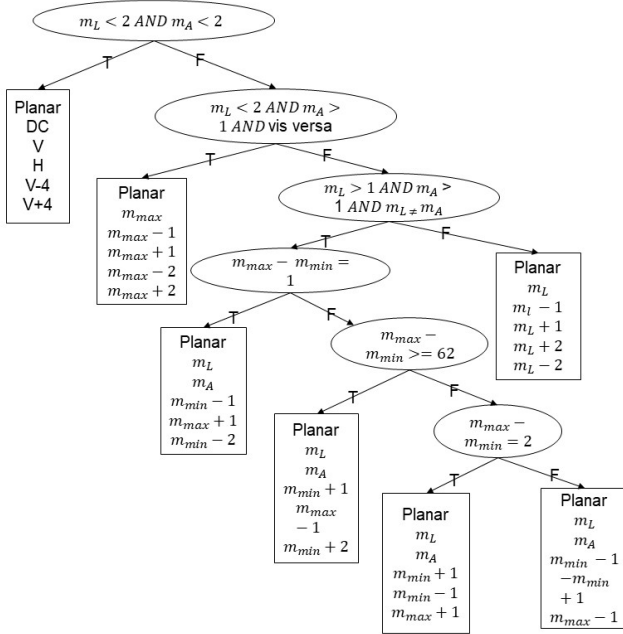


Fig. 4: The 6 MPMs according to the modes of left and above neighbouring blocks m_L and m_A ; m_{max} and m_{min} are the maximum and minimum modes between m_L and m_A , respectively

three step Matrix weighted Intra Prediction (MIP) [7] was integrated in VVC. In fact, when using MIP, the prediction signal is generated in three main steps. Firstly, boundary samples are reduced by averaging according to block size. Then, these averaged samples, are used to generate a reduced prediction signal throw a matrix-vector multiplication, followed by the addition of an offset. Finally, the samples at the remaining positions are interpolated from the reconstructed samples in the matrix-vector multiplication step. This tool uses up to 32 modes [21], which indicate the used matrix and offset vector.

Regarding the interpolation step in intra prediction, VVC uses a 4-tap Gaussian interpolation filter and smoothing filters [21] to generate the prediction signal from reference pixels. Moreover, Position-Dependent Intra Prediction Combination (PDPC) [23] can be applied to reference pixels before carrying out the intra prediction. In fact, depending on the selected intra mode, a combination of filtered and unfiltered reference pixel may be used.

Along with these new tools in partitioning and intra prediction, other tools including Multiple Transform Selection (MTS), Low-Frequency Non-Separable Trans-

form (LFNST), dependent scalar quantization, and Context Adaptive Binary Arithmetic Coding (CABAC) [21], were proposed for transformation, quantization, and entropy coding, respectively.

2.2 The VVC intra mode decision

As in HEVC, the intra mode decision of VVC relies on the Rate-Distortion Optimization (RDO) of the cost J , computed by Eq. (1). This latter measures the encoder's capability to reproduce the source information at a given bitrate [24]. It checks for each coding decision

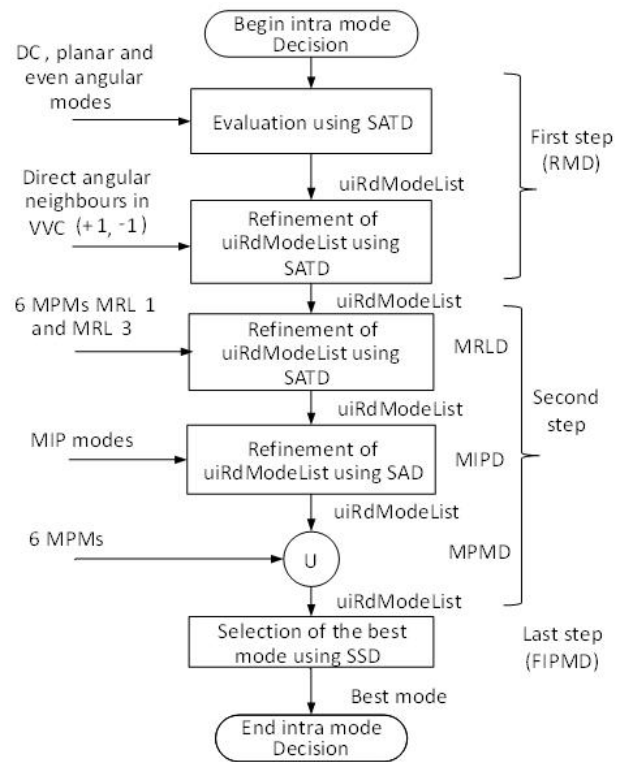


Fig. 5: VVC intra mode decision; uiRdModeList is the list of the retained candidates for the Final Intra Prediction Mode Decision (FIPMD); 6 MPMs MRL 1 and 3 are the lists of 6 MPMs of farther reference lines 1 and 3, respectively

the distortion D and the bit-rate (number of required bits) R , where λ is the Lagrangian multiplier.

$$J = D + \lambda R \quad (1)$$

Due to the large density of IPMs in VVC, the best mode is determined using three main steps, as illustrated in Fig. 5. The first step, so called Rough Mode Decision (RMD), consider only the 67 regular intra modes. It

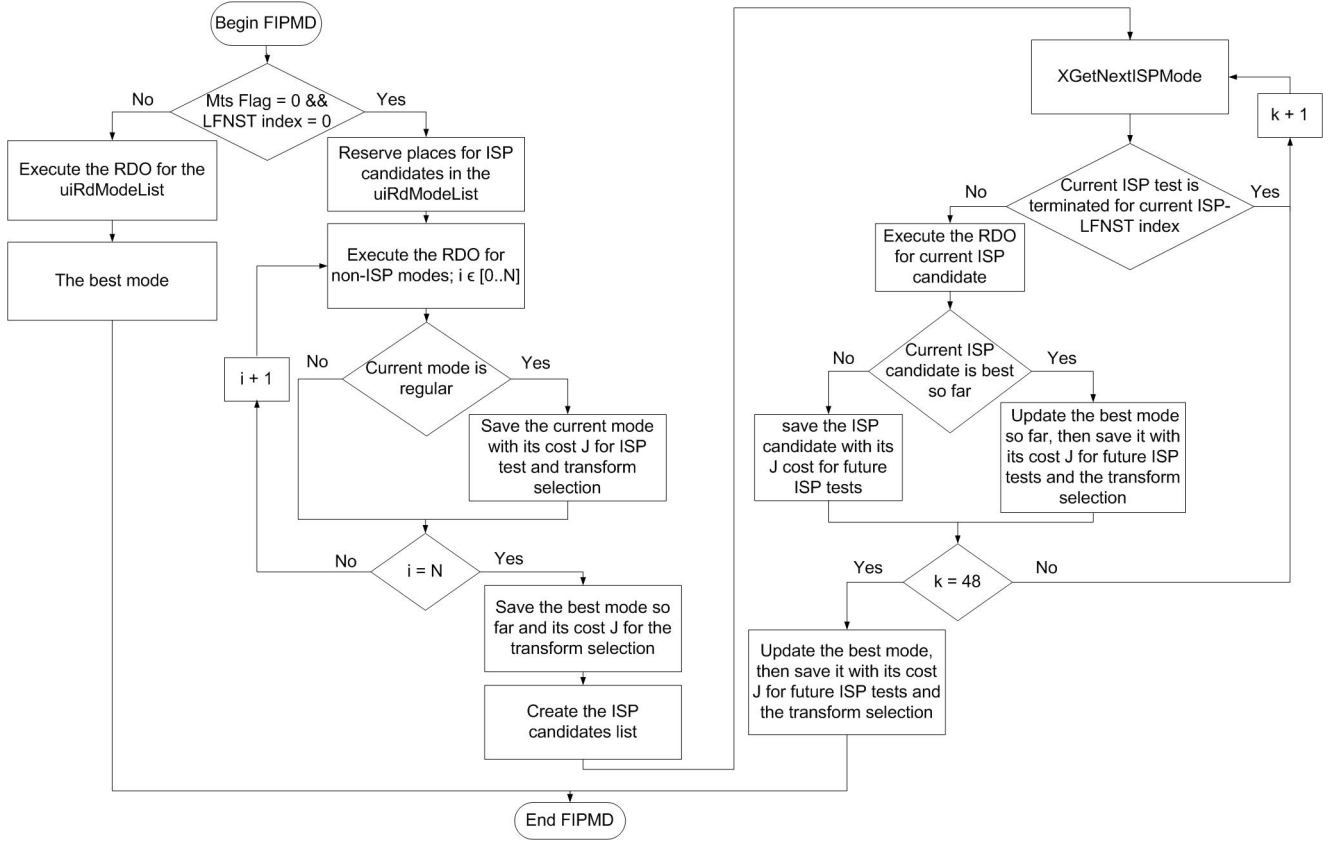


Fig. 6: Final intra Prediction Mode decision; xGetNextISPMODE decides whether to terminate the current ISP test or not

generates a list of N best candidates (uiRdModeList), where $N \in \{2, \dots, 8\}$. First, a test of DC, planar and even angular modes (35 original IPMs of HEVC), is performed. Then, the direct angular neighbours (+1, -1) in the new IPMs of VVC are evaluated. The number of best candidates depends on the block size and whether the MIP is tested for the current block or not. Hence, N is usually between 2 and 3, unless the test of MIP is performed. In this case, the number of best candidates is refined by adding $\log_2(\min(H, W)) - 1$, where H and W are the height and width of the current block, respectively. Additionally, the test of IPMs in RMD is performed using Truncated Binary Coding (TBC) [21] to calculate the bit-rate metric R considered in Eq. (1) and the Sum of Absolute Transform Difference (SATD) as the distortion metric D . This latter is computed by Eq. (2),

$$SATD = \sum_{i,j} |A(i,j)| \quad (2)$$

The matrix A is given by Eq. (3). It corresponds to a matrix multiplication of the residual block with the

Hadamard transform H .

$$A = H \cdot (Y - \hat{Y}) \cdot H^T \quad (3)$$

Hence, Y refers to the original block, $\hat{Y}$ is the predicted block, and H^T denotes the matrix transpose of H .

The second step includes three main decisions: Multi-Reference Lines Decision (MRLD), Matrix weighted Intra Prediction Decision (MIPD), and Most Probable Modes Decision (MPMD). It consists in updating the list of N best candidates using different lists of intra modes. First, the two lists of 6 MPMs: MRL1, and MRL3, where 1 and 3 indicate the reference line indices, are tested using CABAC to infer the number of required bits [21] and SATD as the distortion metric D . Second, the list of MIP modes is tested using TBC and the Sum of Absolute Difference (SAD), defined in Eq. (4). Finally, the list of the 6 regular MPMs is added to the best candidates if they are not already included [21].

$$SAD = \sum_{i,j} |Y(i,j) - \hat{Y}(i,j)| \quad (4)$$

The last step is the Final Intra Prediction Mode Decision (FIPMD). As depicted in Fig. 6, this decision

step may incorporate different tests in which the sum of Squared Difference (SSD), computed by Eq. (5), is used as the distortion metric D . The bit-rate metric R is computed using CABAC. Thus, the best IPM is updated progressively during the FIPMD. According to the MTS flag and the LFNST [21] index, the encoder decides whether to test the ISP for the current block or not. For instance, the ISP cannot be combined with the LFNST or the MTS tools. Hence, when the MTS flag and the LFNST index are both different from zero, only the best IPMs, that were retained following the second step, are tested. Otherwise, the FIPMD can be performed on up to 62 IPMs, considering a maximum of 8 best candidates, 48 ISP modes and 6 MPMs. In this case, the encoder firstly tests the non-ISP modes and saves the costs J of the regular ones. These last will be used later to construct the list of ISP candidates and select the transform core as well. In other words, if the ISP is tested, the VVC encoder creates an ordered list of potential ISP candidates using the best regular mode, non-angular modes (Planar and DC), and the remaining regular modes that were tested during the test of non-ISP modes. It also adds the list of 6 MPMs and the remaining regular modes from the RMD when the ISP is more likely to be selected. Then, the encoder triggers the Intra Sub-Partitions Decision (ISPD) on the different ISP candidates. As a first step, the encoder checks whether the current ISP test can be skipped based on the performance of the previously tested ISP or non-ISP modes. Therefore, if the ISP test has proceeded, then the encoder tests the different ISP partitioning modes for each ISP candidate and save their costs for future ISP tests and transform selection as well. More specifically, the transform selection for blocks using ISP will be performed based on these results instead of the RDO. Unlike, the others steps of intra mode decision, the computation of the cost J is done using the reconstructed block instead of the intra-coded block. Hence, $\hat{Y}$ in Eq. (5) is computed following an overall encoding process. In addition, the number of required bits R is inferred from an entropy coding of the residual block ($Y - \hat{Y}$) instead of an entropy coding of the selected IPM.

$$SSD = \sum_{i,j} |Y(i,j) - \hat{Y}(i,j)|^2 \quad (5)$$

3 Complexity assessment of the VVC intra prediction

This section is devoted to a complexity assessment of the VVC intra prediction. It identifies the most time-

demanding intra coding tools and intra mode decision steps.

3.1 Test conditions and framework

In this work, simulations were conducted on the reference software VVC Test Model (VTM)-8.0 [25] under the Common Test Conditions (CTC) [26]. As shown in Table 1, a set of 26 videos from classes A1, A2, B, C, D, E, and F are used during all simulations. The coding configuration is the All Intra (AI) and the Quantization Parameter (QP) is fixed to 4 values, which are 22,27,32, and 37. Considering that the luma component has the heaviest decision [9] process, all the simulations were carried out only on this component. For the complexity assessment, the number of encoded frames in each video is set to one Group of Pictures (GOP), which is equivalent to one second.

Table 1: Video sequences

Class	Videos	Frame rate	Bit depth
A1-3840 × 2180	<i>Tango2</i>	60	10
	<i>FoodMarket4</i>	60	10
	<i>Campfire</i>	30	10
A2-3840 × 2180	<i>CatRobot1</i>	60	10
	<i>DaylightRoad</i>	60	10
	<i>ParkRunning3</i>	50	10
B-1920 × 1080	<i>MarketPlace</i>	60	10
	<i>RitualDance</i>	60	10
	<i>Cactus</i>	50	8
	<i>BasketBallDrive</i>	50	8
	<i>BQTerrace</i>	60	8
C-832 × 480	<i>RaceHorsesC</i>	30	8
	<i>BasketBallDrill</i>	50	8
	<i>BQMall</i>	60	8
	<i>PartyScene</i>	50	8
D-416 × 240	<i>RaceHorses</i>	30	8
	<i>BQSquare</i>	60	8
	<i>BlowingBubbles</i>	50	8
	<i>BasketBallPass</i>	50	8
E-1280 × 720	<i>FourPeople</i>	60	8
	<i>Johnny</i>	60	8
	<i>KristenAndSara</i>	60	8
F- different resolutions	<i>ArenaOfValor</i>	60	8
	<i>BasketBallDrillText</i>	50	8
	<i>SlideEditing</i>	30	8
	<i>SlideShow</i>	20	8

3.2 Profiling of the intra coding tools

To characterize the effect of the new intra coding tools on the encoder runtime and coding efficiency, the two

averages of runtime difference ΔT and ΔT_I , computed by Eq. (6) and Eq. (7), the Bjontegaard Peak Signal to Noise Ratio (BD-PSNR) and Bjontegaard Bitrate (BD-BR) [27] were measured for all the CTC sequences when disabling some combinations of intra coding tools.

$$\Delta T = \frac{1}{4} \sum_{QP_i=22,27,32,37} \frac{T_o(QP_i) - T_r(QP_i)}{T_o} \quad (6)$$

$$\Delta T_I = \frac{1}{4} \sum_{QP_i=22,27,32,37} \frac{T_{oI}(QP_i) - T_{rI}(QP_i)}{T_{oI}} \quad (7)$$

where T_o and T_r present the total encoder runtimes of the VTM-8.0 anchor and the VTM when disabling some combinations of intra coding tools, respectively. While T_{oI} and T_{rI} are their intra mode decision runtimes. Table 2 reports the obtained results. As shown in this table, the ISP is the most time-demanding intra coding tool. Hence, when this last is disabled, the complexity can shift down by 15% and 20.3% on average for the total encoder runtime and the intra mode decision runtime, respectively. In addition, the ISP has the highest gain in coding efficiency among all the intra coding tools. This is due to the fact that ISP allows further partitioning of intra-predicted blocks into several sub-partitions, which would improve prediction from reference samples. Furthermore, the MIP is less time-demanding than the ISP. Consequently, when these two tools are disabled, the encoder runtimes can be reduced significantly but the loss in coding efficiency would become more remarkable.

In conclusion, the intra coding tools are presenting

Table 2: The effect of the new intra coding tools on encoder runtimes and coding efficiency

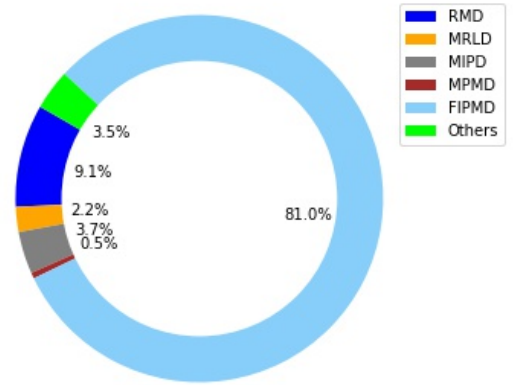
Disabled intra coding tools	BD-BR (%)	ΔT (%)	ΔT_I (%)
MRL	0.43	1.92	2.5
MIP	0.58	12.08	16.7
ISP	0.80	15.04	20.3
$MRL \wedge MIP$	1.05	14.4	19.4
$MRL \wedge ISP$	1.2	16.2	22.0
$MIP \wedge ISP$	1.4	22.4	30.2
$MRL \wedge MIP \wedge ISP$	1.91	28.4	39.3

different tradeoff between complexity and coding efficiency. For example, the ISP considerably improves the coding efficiency, while introducing a huge increase in coding complexity. The MIP and MRL are contributing almost equally to the achieved coding efficiency. However, the MIP has the highest coding complexity.

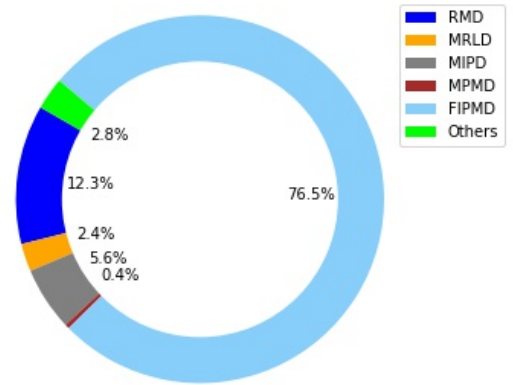
3.3 Profiling of the intra mode decision runtime

To identify the most time-demanding steps of the intra mode decision, a profiling of this decision runtime was performed on the video *Tango2* with all the intra coding tools enabled. Fig. 7 shows the results of this complexity assessment for two QPs.

As depicted in this figure, most of the intra mode decision runtime is related to the FIPMD, with almost 80%. Indeed, beside the additional ISP tests that generates about 3% of additional complexity in the FIPMD, this decision step uses an overall encoding process in order to compute the reconstructed block. This latter may take nearly 70% of the intra mode decision runtime and between 80% to 90% of the FIPMD runtime.



(a) QP=22



(b) QP=37

Fig. 7: Profiling of the intra mode decision runtime in the video *Tango2* for two QPs; (a) QP = 22, (b) QP = 37

4 Complexity reduction opportunities in the VVC intra mode decision

The main focus in this section is to identify the upper bound of complexity reduction in each step of the intra mode decision, while analyzing their impact on the decision runtime when they are disabled.

4.1 Analysis configurations

Since the upper bound of complexity reduction is obtained when the encoder can directly predict the different encoding decisions, an encoding process considering only the best decisions is performed. In this context, the VTM-8.0 encoder [25] is forced to encode only with the best mode(s). For this experiment, two encoding passes were performed. During the first one, the exhaustive Rate-Distortion (RD) search is used to determine the best mode(s). Then, this/these mode(s) are extracted. In the second pass, the extracted mode(s) are then fed back into the encoder in order to discard the complexity and assess the impact of disabling each step of the intra mode decision.

To achieve this, five analysis configurations (C0 to C4), which represent different combinations of intra mode decision steps are defined in Table 3. For each config-

Table 3: Analysis configurations

Intra coding tool	Ref	C0	C1	C2	C3	C4
RMD	E	D	D	D	D	D
MRLD	E	E	D	D	D	D
MIPD	E	E	E	D	D	D
MPMD	E	E	E	E	D	D
FIPMD	E	E	E	E	E	D

uration, the exhaustive RD search can be either enabled (E) or disabled (D). Thus, when this last is disabled at an intra mode decision step, only the best mode(s) related to that decision step are fed back into the encoder. The Ref stands for the VTM-8.0 anchor or the first encoding pass. Whereas, the configurations C0, C1, C2, C3, and C4 refers to the exhaustive RD search being disabled at RMD, at RMD and MRLD, at RMD, MRLD, and MIPD, at RMD, MRLD, MIPD, and MPMD, and finally at all the intra mode decision steps including the FIPMD, respectively. It is worth mentioning that the test of non-ISP modes in the configuration C4 is maintained, as they intervene in the selection of ISP candidates and the transform core as well.

4.2 Analysis of complexity reduction opportunities

The upper bound of complexity reduction in each step of the intra mode decision was evaluated in the intra prediction and the overall encoding process. To this end, the two averages of runtime difference, computed by Eq. (6) and Eq. (7), were used for each analysis configuration. Thus, the average of total encoder runtime difference ΔT between the VTM-8.0 anchor and the VTM when forced to encode only with the best mode(s) is employed for the overall encoding process. Whereas, the average of intra mode decision runtime difference ΔT_I is used for the intra prediction. In this analysis, T_r and T_{rI} present the encoder runtimes of the VTM when forced to encode only with the best mode(s). After computation of these two averages for each CTC video, they are subsequently averaged for all the 26 CTC videos. Hence, Fig. 8 shows these latter as function of the analysis configuration. The blue curve gives the upper bound of complexity reduction for the overall encoding process. Whereas, the orange curve illustrates it for the intra prediction.

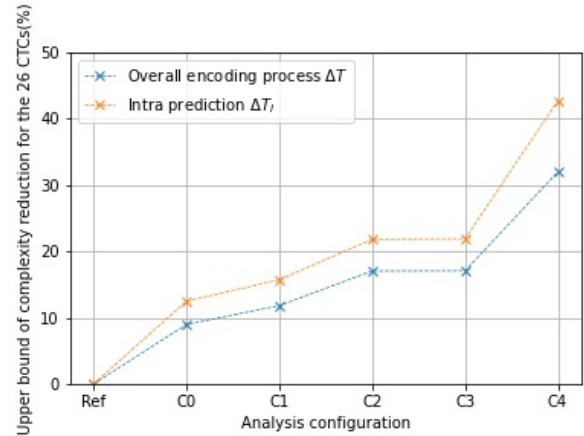


Fig. 8: Upper bound of complexity reduction in the intra mode decision steps for all the 26 CTC videos

As illustrated in this figure, being able to directly predict the best mode(s) at different steps of the intra mode decision could reduce progressively the complexity of both the intra prediction and the overall encoding process. For example, when disabling only the RMD and MRLD in the configuration C1 the upper bound of complexity reduction can reach up to 12% on average in the overall encoding process and 16% in the intra prediction. In addition, most of this reduction, about 43% in average, could be achieved in the configuration C4. More precisely, when directly predicting the best intra mode at FIPMD too. This fact is mainly due to

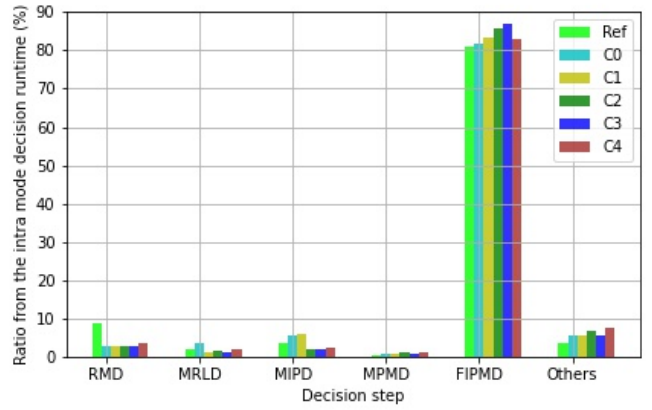
the heaviest computation of the reconstructed block $\hat{Y}$, as explained in section 3. However, this upper bound of complexity reduction could be less significant compared to the VTM-3.0 [11], and this is directly related to the introduction of MIP and ISP tools which generated additional complexity, as shown in Section 3. Moreover, the MPMD introduced almost no difference in the total encoder runtime. Indeed, the MPMD consists in adding modes without any further computation. Thus, one can observe that on average only 0.03% of additional reduction can be introduced in the total encoder runtime of the configuration C2 when compared to the configuration C3. It can also be noticed, from Fig. 8, that all configurations showed a higher upper bound of complexity reduction in the intra prediction than the overall encoding process. This is obvious because the overall encoding process still uses the exhaustive RD search at the partitioning level, which is very time-demanding compared to the intra prediction as explained in [9] and [11]. This fact may also influence the upper bound of complexity reduction in the intra prediction since this step is involved in each RD iteration of the partitioning decision. Thus, targeting simultaneously several decision levels in the exhaustive RD search would lead to higher complexity reduction.

To conclude, a direct prediction of the best mode at the FIPMD could contribute most to the complexity reduction of the VVC encoder. Hence, the upper bound of complexity reduction can reach up to 43% on average in the intra prediction and 32% in the overall encoding process. Moreover, the MIPD and MRLD could introduce additional reductions when they are disabled. These reductions may achieve on average 12% in the intra prediction and 9% in the overall encoding process.

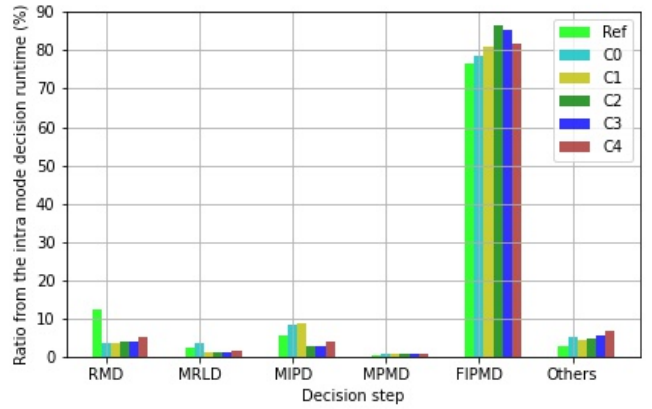
4.3 the impact of disabling each step of the intra mode decision

In Fig. 9, a profiling of the intra mode decision runtime within the different analysis configurations is conducted on the video *Tango2*. Indeed, during the second encoding pass, the impact of disabling each intra mode decision steps is studied.

When analyzing the results of this complexity assessment, it reveals that a direct prediction of the best mode(s) at any step of the intra mode decision could reduce its impact on this decision runtime. For example, when feeding back the best regular modes for the RMD in the configuration C0, its ratio from the intra mode decision runtime decreased significantly. Thus, about 8% of decrease is observed for this decision step. The only exception is presented for the FIPMD as it presents



(a) QP = 22



(b) QP = 37

Fig. 9: Profiling the intra mode decision runtime when disabling each decision step in the video *Tango2* for two QPs; (a) QP = 22, (b) QP = 37

the most time-demanding step of the intra mode decision. Hence, disabling it in configuration C4 could enable a significant upper bound of complexity reduction, and thus making its impact on the intra mode decision runtime more noticeable. Table 4 illustrates a profiling of the FIPMD runtime when the ISP is tested or not tested in the VTM-8.0 anchor, denoted as Ref, and in the configuration C4. It displays also its ratio from the intra mode decision runtime in order to provide a clear understanding of its impact on the intra mode decision runtime for both cases.

As it can be observed in this table, when ISP is tested, the FIPMD impact is more visible in configuration C4 than in Ref. This is due to the fact that the test of non-ISP modes is maintained. More specifically, its ratio from the FIPMD runtime has increased. In this case, the ratio of the ISPD from the FIPMD runtime, even when testing only the best ISP mode, is going to be im-

Table 4: Profiling of FIPMD runtime when ISP is tested or not tested in the video *Tango2* for two QPs

QP	ISP	Ref		C4	
		intra mode decision runtime (%)	FIPMD runtime (%)	intra mode decision runtime (%)	FIPMD runtime (%)
22	Tested	43.2	53.4	57.0	68.6
	Not tested	37.7	46.6	26.1	31.4
37	Tested	43.1	56.3	60.2	73.8
	Not tested	33.4	43.7	21.4	26.2

portant in comparison with the achieved upper bound of complexity reduction in this configuration. The same behavior is observed for the enabled decision steps and the different initialization tasks, which are becoming heavier through the simulations. For the MPMD, its ratio from the intra mode decision runtime stabilized at around 1%. In fact, this decision step doesn't affect much the intra mode decision runtime, as explained in Section 3.

5 Statistical analysis of selected IPMs

For a better understanding of the most time-demanding step of the intra mode decision, an in-depth statistical analysis [28] of the FIPMD according to the intra coding tool is provided in this section. It describes the ratio of selection for each intra coding tool and its most selected IPMs according to the MT depth and video texture characteristics. In order to presents only the most selected IPMs a threshold was set to 1.5%. Therefore, each IPM that has been selected less than this threshold is not presented.

In this section, the average ratio of selection of each IPMs for all CTC sequences is presented. The results are illustrated only for the QP = 22. In addition, this statistical analysis was conducted using the default parameters of the QTMT structure, given in Table 5, where the MT depth is the sum of the two depths of the partitioning trees of the MT structure. Its maximum value is equal to 3 unless the size of the current MT leaf node exceeds the frame boundaries. Hence, MT0 presents 64×64 and the maximum MT depth may reach 5 in some cases. Fig. 10 illustrates the average ratio of selection of each intra coding tool according to the MT depth.

As it can be seen from this figure, the FIPMD tends usually to select regular modes regardless of the MT

Table 5: Default parameters of the QTMT structure

QTMT Parameter	Value
Size of the CTU	128×128
MinQTsize	16×16
MaxBTsize and maxTTsize	64×64
MinBTsize and MinTTsize	4
MaxMTDepth	4

depth. Moreover, the MRL and MIP tools are often chosen when regular modes are not selected. As a con-

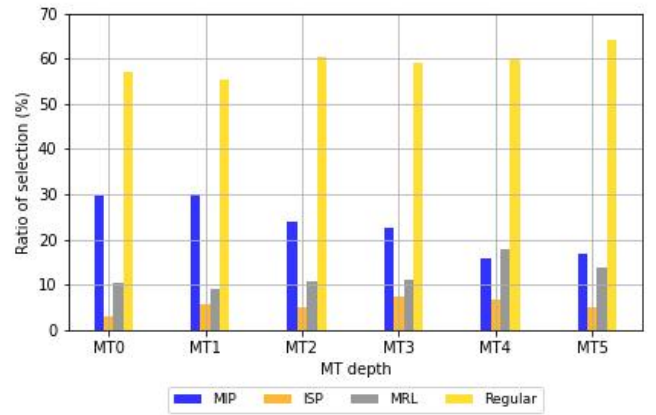


Fig. 10: the average ratio of selection of each intra coding tool according to the MT depth for QP = 22

sequence, for each intra coding tool, the FIPMD is analyzed.

Figures 11, 12 and 13 display the most selected IPMs when regular, ISP, and MIP tools are selected, respectively. From figures 11 and 12, it reveals that when regular or ISP are used, several modes from the list of the 6 MPMs, depicted in Fig. 4, are the most selected for all MT depths. In fact, the use of the 6 MPMs helps in reducing the signaling overhead of the IPMs in the bit-stream [29]. Besides, the IPMs around the bottom-left direction (mode 2) are not commonly used. However, they are more likely to be selected if the MIP is used, as described in Fig. 13. This is a result of the shifting operation that has been applied to the other half of the MIP modes during the MIPD. As such, a high likelihood of selection would be granted to these directions. Furthermore, the ISP allows the selection of various directions of intra prediction and this is due to the use of the 6 MPMs along with regular modes. This behavior may result in improving objective visual quality and coding efficiency. Fig. 14 highlights the most selected

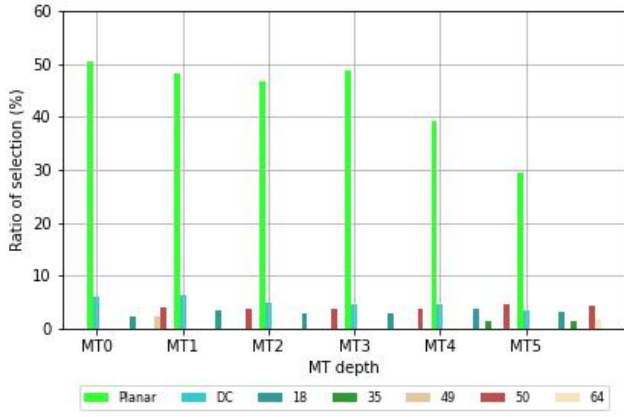


Fig. 11: Most selected IPMs according to the MT depth when regular modes are selected and $QP = 22$

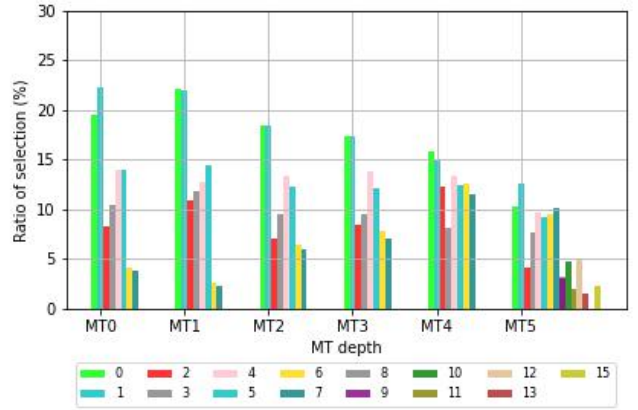


Fig. 13: Most selected IPMs according to the MT depth when MIP is selected and $QP = 22$

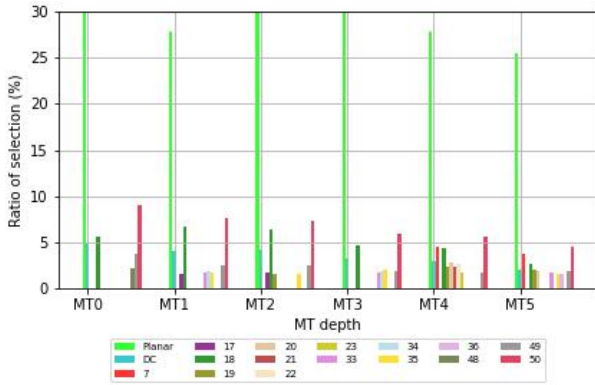


Fig. 12: Most selected IPMs according to the MT depth when ISP is selected and $QP = 22$

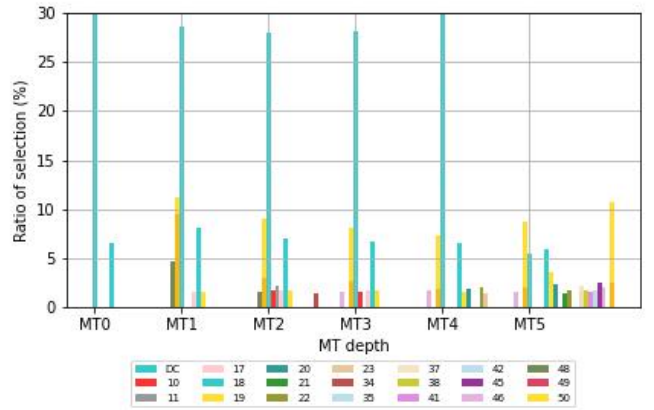


Fig. 14: Most selected mode according to the MT depth when MRL is selected and $QP = 22$

IPMs when the MRL is used. As shown in this figure, the 6 MPMs other than the Planar are often used for this tool. This fact is explained by the incompatibility of Planar mode with the MRL [21].

All in all, though, the new intra coding tools have shown different FIPMD behavior as they are using various sets of IPMs. As an example, the MIP presents a simplified neural network-based intra prediction, where the best mode is selected from a set of 32 trained modes. The MRL is also a good example since it considers only the 6 MPMs modes. The only exception is shown for the ISP as it considers a list of IPMs created based on the regular modes. As such the FIPMD, when ISP is used, may depend strongly on these modes. When studying also the results of the statistical analysis for all the intra coding tools it can be noticed that additional IPMs, such as 22 and 23, are slightly used along with the most selected IPMs and this is due to the texture richness of same CTC videos and the diversity of their contents.

6 conclusion

In this paper, a complexity assessment of the intra mode decision is conducted. It first identifies the most time-demanding intra coding tools and intra mode decision steps. Then, the upper bound of complexity reduction in each of these steps and tools is analyzed while discussing the impact of disabling the RD search on this decision runtime. Finally, an in-depth statistical analysis of the most time-demanding step of the intra mode decision, which is the FIPMD, is given.

Results showed that the intra coding tools and the intra mode decision steps are presenting varying degrees of complexity. Thus, revealing several complexity reduction opportunities. For instance, the complexity of the VVC can be reduced by up to 43%. Also, the FIPMD tends usually to select the 6 MPMs. However, this selection may depend on the used intra coding tool and video texture characteristics.

This assessment may have potential limitations. In fact, it has been conducted on the CTC sequences and the results have been reported regardless of any complexity reduction technique. Hence, these latter may be subject to variation depending on the used complexity reduction technique and video sequences. Nevertheless, this assessment underlined the potential complexity reduction opportunities in the intra mode decision, which can be exploited to speed up the RD search. In addition, the existing correlations between video content and the RD search were discussed accordingly, which would help the video coding community in developing robust decision algorithms, like machine learning, in order to alleviate the complexity of the VVC and allow real-time VVC encoding.

Acknowledgements This work is supported by the France Campus, and within a co-supervised thesis between the Institute of Electronics and Telecommunications of Rennes, France (IETR), and the Laboratory of Electronics and Information Technology (LETI) of Sfax, Tunisia.

Conflict of interest

This study was funded by France Campus with the PHC Maghreb project Eco-VVC (grant number 45988WG). N. Zouidi, A. Kessentini, W. Hamidouche N. Masmoudi and D. Menard declare that they have no conflict of interest.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Cisco: Cisco visual networking index: Forecast and trends, 2017 to 2022 (2018). [online] available at <https://cloud.report/whitepapers/cisco-visual-networking-index-forecast-and-trends-2017-2022/9017>
2. Akramullah, S.: Digital video concepts, methods, and metrics: quality, compression, performance, and power trade-off analysis. Springer Nature (2014)
3. JVET: AHG report: Test model software development (AHG3). Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-B1010 (19th Meeting, teleconference, July, JVET-S0003 (2020))
4. Li, X.e.a.: Multi-Type-Tree, document JVET-D0117. 4th JVET meeting, Chengdu, CN (October 2016)
5. De-Luxán-Hernández, S., George, V., Ma, J., Nguyen, T., Schwarz, H., Marpe, D., Wiegand, T.: An intra subpartition coding mode for vvc. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 1203–1207. IEEE (2019)
6. Li, J., Li, B., Xu, J., Xiong, R.: Intra prediction using multiple reference lines for video coding. In: 2017 Data Compression Conference (DCC), pp. 221–230. IEEE (2017)
7. Schäfer, M., Stallenberger, B., Pfaff, J., Helle, P., Schwarz, H., Marpe, D., Wiegand, T.: An affine-linear intra prediction with complexity constraints. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 1089–1093. IEEE (2019)
8. JVET: AHG report: Test model software development (AHG3). Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 (18th Meeting: by teleconference, 15 to 24 April 2020, JVET-R0003)
9. Saldanha, M., Sanchez, G., Marcon, C., Agostini, L.: Complexity analysis of vvc intra coding. In: 2020 IEEE International Conference on Image Processing (ICIP), pp. 3119–3123. IEEE (2020)
10. Pakdaman, F., Adelimanesh, M.A., Gabbouj, M., Hashemi, M.R.: Complexity analysis of next-generation vvc encoding and decoding. arXiv preprint arXiv:2005.10801 (2020)
11. Tissier, A., Mercat, A., Amestoy, T., Hamidouche, W., Vanne, J., Menard, D.: Complexity reduction opportunities in the future vvc intra encoder. In: 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. IEEE (2019)
12. Cao, J., Tang, N., Wang, J., Liang, F.: Texture-based fast cu size decision and intra mode decision algorithm for vvc. In: International Conference on Multimedia Modeling, pp. 739–751. Springer (2020)
13. Chen, Y., Yu, L., Wang, H., Li, T., Wang, S.: A novel fast intra mode decision for versatile video coding. Journal of Visual Communication and Image Representation **71**, 102849 (2020)
14. Yang, H., Shen, L., Dong, X., Ding, Q., An, P., Jiang, G.: Low complexity ctu partition structure decision and fast intra mode decision for versatile video coding. IEEE Transactions on Circuits and Systems for Video Technology (2019)
15. Zouidi, N., Kessentini, A., Belgith, F., Masmoudi, N.: Statistical analysis of qtbt structure of versatile video coding encoded videos. In: 2019 16th International Multi-Conference on Systems, Signals & Devices (SSD), pp. 18–21. IEEE (2019)
16. Zouidi, N., Belghith, F., Kessentini, A., Masmoudi, N.: Fast intra prediction decision algorithm for the qtbt structure. In: 2019 IEEE International Conference on Design & Test of Integrated Micro & Nano-Systems (DTS), pp. 1–6. IEEE (2019)
17. Siqueira, I., Correa, G., Grellert, M.: Complexity and coding efficiency assessment of the versatile video coding standard. In: 2021 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–5. IEEE (2021)
18. Mercat, A., Mäkinen, A., Sainio, J., Lemmetti, A., Viitanen, M., Vanne, J.: Comparative rate-distortion-complexity analysis of vvc and hevcd video codecs. IEEE Access **9**, 67813–67828 (2021)
19. Saldanha, M., Sanchez, G., Marcon, C., Agostini, L.: Performance analysis of vvc intra coding. Journal of Visual Communication and Image Representation **79**, 103202 (2021)
20. Pfaff, J., Filippov, A., Liu, S., Zhao, X., Chen, J., De-Luxán-Hernández, S., Wiegand, T., Ruffitskiy, V., Ramasubramanian, A.K., Van der Auwera, G.: Intra prediction and mode coding in vvc. IEEE Transactions on Circuits and Systems for Video Technology **31**(10), 3834 to 3847 (2021)

21. JVET: Algorithm description for Versatile Video Coding and Test Model 8. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 17th Meeting: Brussels, BE (17 January 2020)
22. Lainema, J., Bossen, F., Han, W.J., Min, J., Ugur, K.: Intra coding of the hevc standard. *IEEE transactions on circuits and systems for video technology* **22**(12), 1792–1801 (2012)
23. Said, A., Zhao, X., Karczewicz, M., Chen, J., Zou, F.: Position dependent prediction combination for intra-frame video coding. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 534–538. IEEE (2016)
24. Sullivan, G.J., Wiegand, T.: Rate-distortion optimization for video compression. *IEEE signal processing magazine* **15**(6), 74–90 (1998)
25. JVET: Vvcsoftware_vtm 8.0 reference software (2018). [online] available at https://vcgit.hhi.fraunhofer.de/jvet/VVCSSoftware_VTM/
26. JVET: JVET common test conditions and software reference configurations for SDR videos. Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-B1010 (14th Meeting: Geneva, CH, 19 to 27 Mar. 2019)
27. Bjontegaard, G.: Document vceg-m33: Calculation of average psnr differences between rd-curves. In: ITU-T VCEG Meeting, Austin, Texas, Technical Report (2001)
28. Zouidi, N., Kessentini, A., Belghith, F., Masmoudi, N.: Statistical analysis of the qtmt structure: intra mode decision. In: 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS), pp. 180–185. IEEE (2020). DOI 10.1109/IPAS50080.2020.9334954
29. REUZE, K., PHILIPPE, P., DEFORGES, O., HAMIDOUCHE, W.: Enhanced intra prediction modes signalling in hevc. In: 2016 Picture Coding Symposium, PCS 2016. Nuremberg, Germany (2016). DOI 10.1109/PCS.2016.7906387. URL <https://hal-univ-rennes1.archives-ouvertes.fr/hal-01622528>