Complexity Reduction Opportunities in the Future VVC Intra Encoder

A. Tissier[†], A. Mercat[⋆], T. Amestoy[‡†], W. Hamidouche[†], J. Vanne[⋆] and D. Menard[†]

[†] Univ Rennes, INSA Rennes, CNRS, IETR - UMR 6164, Rennes, France.

[⋆] Laboratory of Pervasive Computing, Tampere University of Technology, Tampere, Finland.

[‡]Thales SIX GTS France, HTE/STR/MMP Gennevilliers, France.

Emails: alexandre.tissier2@insa-rennes.fr, alexandre.mercat@tuni.fi, thomas.amestoy@insa-rennes.fr,
wassim.hamidouche@insa-rennes.fr, jarno.vanne@tuni.fi, daniel.menard@insa-rennes.fr

Abstract—The Joint Video Expert Team (JVET) is developing the next-generation video coding standard called Versatile Video Coding (VVC) and their ultimate goal is to double the coding efficiency over the current state-of-the-art standard HEVC without letting complexity get out of hand. This work addresses the complexity of the VVC reference encoder called VVC Test Model (VTM) under All Intra coding configuration. The VTM3.0 is able to improve intra coding efficiency by 21% over the latest HEVC reference encoder HM16.19. This coding gain primarily stems from three new coding tools. First, the HEVC Quad-Tree (QT) structure extension with Multi-Type Tree (MTT) partitioning. Second, the duplication of intra prediction modes from 35 to 67. And third, the Multiple Transform Selection (MTS) scheme with two new discrete cosine/sine transforms (DCT-VIII and DST-VII). However, these new tools also play an integral part in making VTM intra encoding around 20 times as complex as that of HM. The purpose of this work is to analyze these tools individually and specify theoretical upper limits for their complexity reduction. According to our evaluations, the complexity reduction opportunity of block partitioning is up to 97%, i.e., the encoding complexity would drop down to 3% for the same coding efficiency if optimal block partitioning could be directly predicted. The respective percentages for intra mode reduction and MTS optimization are 65% and 55%. We believe these results motivate VVC codec designers to develop techniques that are able to take most out of these opportunities.

Index Terms—Versatile Video Coding (VVC), Complexity reduction, CTU partitioning, Multi-Type Tree (MTT), Intra mode prediction, Multiple Transform Selection (MTS)

I. INTRODUCTION

IP video traffic is estimated to account for up to 82% of the global IP traffic by 2022 [1]. Considering this evolution through the emerging video formats like 4K Ultra High Definition (UHD) and 360-degree videos, the coding performance of the current High Efficiency Video Coding (HEVC) standard [2] needs to be further enhanced to satisfy the requirements of future video streaming and storage. These new challenges motivated the International Telecommunication Union (ITU) and ISO/IEC Moving Picture Experts Group (MPEG) to form the Joint Video Expert Team (JVET) that is currently developing a new video coding standard called Versatile Video Coding (VVC) [3].

JVET is investigating many new coding tools in All Intra (AI) configuration such as the MTT partitioning, 32 new angular intra mode predictions, and 2 new transforms. These new tools have been integrated into the VVC reference software called VVC Test Model (VTM). First, VTM extends the Quad-Tree (QT) block partitioning scheme of HEVC by adding the nested recursive MTT partitioning with additional Coding Unit (CU) types. This new block partitioning scheme forms the basis of the VTM encoding process with the highest coding gain. So, VTM allows five different splits with QT, Binary-Tree (BT) and Ternary-Tree (TT). BT and TT include both horizontal or vertical partitioning. Secondly, the number of intra prediction modes is extended from that of 35 in HEVC to 67 in VTM3.0 to better leverage spatial redundancy of the reconstructed neighboring blocks. Thirdly, VTM introduces the MTS process that tests different core transforms and selects the one with the best coding efficiency. In addition to the Discrete Cosine Transform (DCT)-II adopted from HEVC, two new transforms, DCT-VIII and Discrete Sine Transform (DST)-VII, are included in MTS.

The new coding tools of VTM3.0 improve average coding efficiency by 21.08% [4] over that of HEVC test Model (HM) 16.19, in AI coding configurations. These improvements come at the expense of 1919% [4] encoding complexity. Reduction of encoding complexity was already a hot topic during HEVC standardization. Considering the aforementioned complexity increase of VTM3.0, encoding complexity reduction is supposed to remain an active research field during the VVC standardization.

In this context, we propose a hierarchical characterization and evaluation of the VVC encoding complexity with the VTM encoder. We first present an overview of the impact of encoding parameters on the encoding complexity. The investigated parameters include spatial resolution and the Quantization Parameter (QP). Then, we define and evaluate the complexity reduction opportunities offered by the three algorithmic encoding levels: Coding Tree Unit (CTU) partitioning, intra mode prediction and MTS process.

The remainder of this paper is organized as follow. Section II presents an overview of the VVC encoding tools and the

state-of-the-art of complexity reduction techniques. Section III details the experimental setup and evaluates the encoding complexity with different encoding parameters. Section IV presents the identified complexity reduction opportunities and analyses their impact on the VVC encoding process. Finally, Section V concludes the paper.

II. RELATED WORKS

This section first presents an overview of the encoding process, from HEVC to VVC. As this work is dedicated to AI configuration, the focus is on Intra coding tools. State-of-the-art complexity reduction techniques are then introduced.

A. From HEVC to VVC encoding process overview

1) HEVC: In HEVC [2], adaptive block partitioning is the primary factor for coding gain. HEVC encoder splits each frame into equally-sized blocks called CTUs. Each CTU is further divided into CUs, which are recursively split into sub-CUs following a QT partitioning. CUs may be split into Prediction Units (PUs) of smaller size, on which the prediction is performed. HEVC enables 35 intra prediction modes including DC, planar, and 33 angular modes. After prediction, the transform and quantization steps are applied on the residual blocks, called Transform Units (TUs). HEVC includes DCT-II and DST-VII transforms, out of which DST-VII is only applied on luma 4×4 blocks.

2) VVC: Versatile Video Coding (VVC) is defined to be the next generation video coding standard, which has as primary target to provide a significant performance improvement over HEVC. VTM is the VVC reference software used by JVET to evaluate the performance of the new encoding tools. As VVC is expected to be standardized by 2020, the following description of the encoding process is focused on VTM3.0.

In addition to the recursive QT partitioning of HEVC, VTM integrates a nested recursive MTT partitioning, i.e., BT and TT splits. Fig. 1 illustrates all available splits in VTM for a $4N \times 4N$ CU. The BT partitioning consists of symmetric horizontal splitting (BT-H) and symmetric vertical splitting (BT-V). The TT partitioning allows horizontal triple-tree splitting (TT-H) and vertical triple-tree splitting (TT-V) corresponding to split the CU in three blocks so that size of the middle block is half of the CU size. In the AI configuration file defined in the Common Test Conditions (CTC), BT and TT are available on CUs with sizes lower or equal than 64×64 and 32×32, respectively. Moreover, the partitioning process includes restrictions to avoid generating the same CU with different succession of split. For example, BT partitioning is not allowed on the middle CU of a TT split in the same direction. Once BT or TT split is performed on a CU, QT split is not allowed any more on its sub-CUs.

Fig. 2(a) presents an example of a CTU split into several CUs after performing the VTM CTU partitioning with QT and MTT splits and Fig. 2(b) shows a part of CTU partitioning corresponding tree representation. Dark, green and blue lines represent QT split, BT split, and TT split, respectively. Yellow
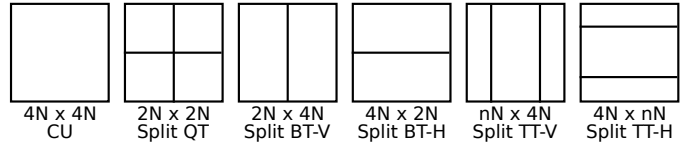

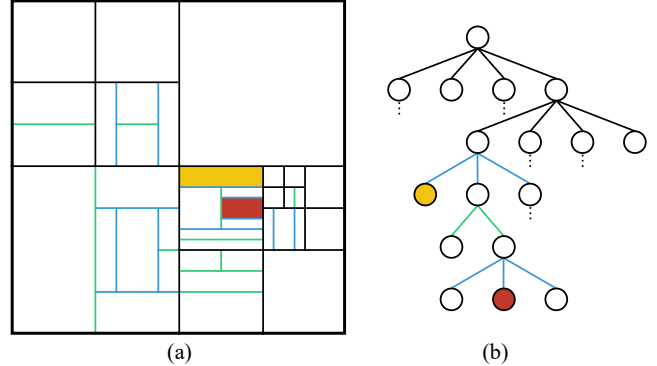Fig. 1. Available split included in VTM of a $4N \times 4N$ CU.

Fig. 2. Example of a CTU partitioning (a) with a part of its corresponding tree representation (b).

and red background show two examples of the corresponding CUs on the CTU partitioning and corresponding tree.

To the best of our knowledge, no performance comparison between CTU partitioning of VTM and HM has been made in AI configuration. For information, adding BT and TT to HM16.14 encoder in Random Access (RA) configuration increases the Bjøntegaard Delta Bit Rate (BD-BR) performance by -14.42% at the expanse of encoding complexity increase of 183% [5].

Intra mode prediction in VTM is similar to that of HEVC. Intra prediction is made from reconstructed pixel samples of neighboring PUs. Angular modes in VTM are extended to 65. As the VTM allows rectangular shape for a PU, wide-angle modes replace some angular modes according to the dimensions (width / height) of the PU [6]. Increasing the number of intra modes from 35 to 67 in VVC provides, on average, -0.51% BD-BR saving for 108% encoding complexity increase in AI configuration [7].

For the transform step, VTM introduces the MTS process that enables two new transforms including DST-VII and DCT-VIII. MTS tests a set of transforms and selects the one that achieves the best encoding efficiency. For luma component, MTS works on TU of size equal or lower than 32×32, otherwise DCT-II is selected. For chroma component, only DCT-II is considered. In VTM3.0, MTS provides BD-BR saving of up to -2.8% for an encoding complexity increase of 238% in AI configuration [7].

To achieve the best RD performance, the encoder performs an exhaustive search process, named Rate-Distortion Optimization (RDO), testing all possible combinations of CTU partitioning structures, intra prediction modes, and transforms. The RDO process minimizes the cost J, called Rate-Distortion (RD)-cost, defined as

$$J = D + \lambda \cdot R \qquad (1)$$

where D is the distortion, R the bit-rate, and λ the Lagrangian weighting factor, which depends on the QP.

B. State-of-the-Art of complexity reduction techniques

The new encoding tools presented in Section II-A1, especially adaptive block partitioning, have increased encoding complexity compared to previous standards. Many methods have already been proposed to reduce the encoding complexity of HEVC. To evaluate their results, our previous work [8] presents an analysis of complexity reduction opportunities of an HEVC real-time intra encoder *Kvazaar*. This study shows that the CTU partitioning process has a potential of complexity reduction of up to 78.1% whereas the intra mode prediction offers at best 30% complexity reduction. This section presents the state-of-the-art of complexity reduction focused on CTU partitioning and intra mode prediction on HM and Joint Exploration Model (JEM) encoders.

The following methods reduce the complexity of QT partitioning. The authors in [9] and [10] predict split decision with machine learning techniques on HM12.0. The first method [9] uses multiple Convolutional Neural Networks (CNN) depending on the CU size to predict CU and PU split decision. It achieves, on average, 61.1% complexity reduction in AI configuration for a BD-BR loss of +2.67%. The second method [10] presents three sets of decision trees that early terminate the RDO process of CU, PU, and TU. This technique enables 65% complexity reduction for +1.36% BD-BR loss in RA configuration. Authors in [11] propose a solution combining CU depth pre-selection, early CU and PU terminations, and fast TU tree decision based on a RD complexity optimization formula. The complexity gain of the four combined propositions on HM16.7 in RA configuration is 46% to 70% for +0.48% to +2.36% BD-BR loss, respectively.

QTBT partitioning was introduced in HM13.0-QTBT which has considerably increased the encoding complexity compared with QT partitioning. The method proposed in [12] combines two solutions to reduce the complexity of HM13.0-QTBT. The first solution dynamically sets partition parameters at CTU level according to splitting information of neighboring blocks. The second solution is a joint-classifier that early terminates the CTU partitioning process. With +1.34% BD-BR loss, this method obtains a complexity reduction of 67.6% in AI configuration. In [13], authors propose early termination algorithm based on their proposed RD model. This RD model is based on motion divergence field and predicts the RD cost at each partition pattern without full RDO process. This method was implemented on JEM7.0 and it enabled 50.60% complexity reduction for +1.23% BD-BR loss in RA configuration. The method presented in [14] uses CNN to predict the depth range of 32×32 block size. The use of CNN come with an overhead of 3.36% and introduces +0.69% BD-BR loss for complexity reduction of 42.33% on JEM3.1.

In order to reduce the complexity of intra mode prediction, authors in [15] propose a progressive rough mode search based on the Hadamard cost to selectively evaluate potential prediction modes. Furthermore, this progressive rough mode search is completed by early rate-distortion optimized quantization skip to further reduce the set of tested intra modes. The combination of these two methods implemented in HM4.0 offers a complexity gain of 45% for +0.8% BD-BR loss. Machine learning solutions [16] [17] limit the number of intra prediction modes. These solutions reduce the encoding complexity of 23.61% and 18.3% for +0.03% and +0.5% BD-BR loss on HM16, respectively.

As presented in Section II-A, MTS process included since the first version of VTM brings a significant complexity increase. To the best of our knowledge, no solution have been proposed to reduce the complexity of the MTS process. Finally, a first method [18] implemented in VTM2.0 proposes early termination of QT plus MTT partitioning and intra mode prediction to skip unnecessary partition and intra modes. The proposed method obtains 63% complexity reduction at the expense of +1.93% BD-BR loss.

III. COMPLEXITY ANALYSIS OF VTM INTRA ENCODER

Complexity reduction is a key aspect on VTM. Complementary to the state-of-the-art complexity reduction techniques introduced above, this paper characterizes and evaluates the encoding complexity of the VTM encoder. This work takes a hierarchical approach, from the encoding parameter level to the encoding tool level including 1) CTU partitioning; 2) intra mode prediction; and 3) MTS process.

A. Experimental setup

Our experiments are performed on VTM3.0 under AI coding configuration and the JVET CTC [19]. CTC are specified by JVET to conduct experiments in a well-defined environment. Our test set contains 22 sequences, which belong to the following 5 CTC classes: A (3840×2160), B (1920×1080), C (832×480), D (416×240), and E (1280×720). Class A is further divided in two sub-classes: A1 and A2. These classes feature different frame rates, bit depths, motions, textures, and spatial resolutions. Each video sequence is encoded with four QP values of 22, 27, 32, and 37. All test are carried out sequentially on Intel Xeon E5-2603 v4 processors running the Ubuntu 16.04.5 operating system. The following sections examine the impact of two encoding parameters: video resolution and QP.

B. Complexity impact on video resolution

The VTM encoder applies the same coding process on each CTU, so its complexity is directly linked to the number of CTUs and consequently to the spatial resolution and frame rate of the encoded sequence.

Fig. 3 displays the average encoding time per frame in seconds according to the number of pixels per frame. The average encoding time per frame is plotted in blue for the different CTC classes (detailed in Section III-A). Blue area represents the standard deviation of the results across the different sequences of the same class. In addition, the red line shows the average encoding time per frame for the *Tango2*
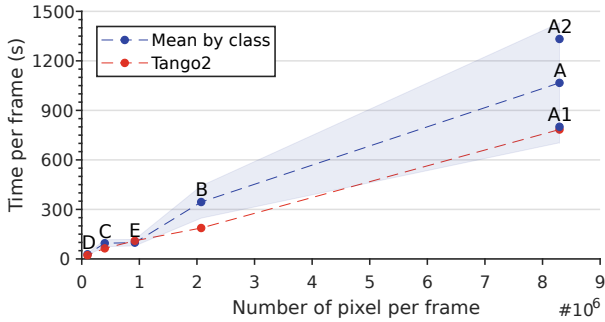
Fig. 3. Average encoding time per frame in seconds as a function of the number of pixels per frame for each class and Tango2 downscaled encodings.

Fig. 4. Average encoding time per CTU as a function of the QP, classes A-E.

sequence which was downscaled by the FFmpeg bilinear filter in all CTC resolutions. *Tango2* sequence was chosen since its average encoding time per frame was close to the average of its class A1.

Fig. 3 shows that the complexity of the VTM encoder increases with the number of pixels per frame except for class E, where the sequences have specific screen content properties such as fixed unified background that explains the irregularity. The encoding time increases linearly with the number of pixels for the same video content, as shown with the downscaled *Tango2* sequences (in red in Fig. 3). However, the standard deviation of the mean varies between 19% and 34%, which highlights that the encoding complexity is highly linked to the video content.

The VTM encoder includes content-dependent early termination mechanisms in the RDO process. One of these methods compares the RD-cost of the unsplit CU with the accumulated RD-cost of sub-CUs to terminate the partitioning process when the accumulated cost becomes higher than the one of the unsplit CU. These complexity reduction methods explain the non-stability of encoding complexity across different sequences of varying video content. High complexity difference in the same resolution is further shown by the two sub-classes A, with the difference of 532s between average time per frame for class A1 and A2.

C. Complexity impact on QP

As explained in Section II-A, the QP is used both in the RD-cost calculation and in the quantification step of the encoding process. Fig. 4 illustrates the average encoding time per CTU for each class as a function of QP. The results show that the encoding complexity decreases as QP increases for all classes. For example, for class B, the encoding time per CTU is 4.75s at QP 22 and 1.02s at QP 37. This is mainly due to the fact that an encoding with higher QP quantizes data more aggressively, leading to a larger number of zero coefficients after quantization.

The results of Fig. 4 also show that the complexity reduction due to QP does not behave consistently between the classes. When dividing the time per CTU at QP 22 by the time at QP 37, the factors are 7.61 and 2.24 for class A and D, respectively. The higher the resolution, the greater the
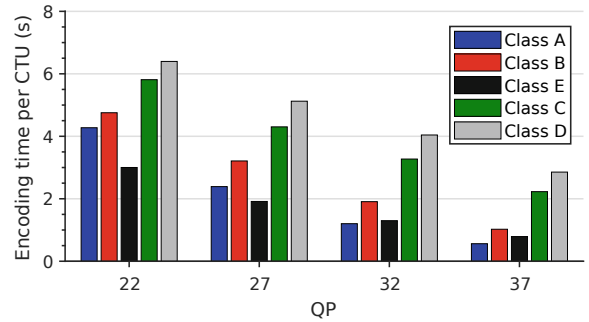
impact of QP on encoding times per CTU. This can be explained by the complexity reduction techniques mentioned in the Section III-C. Indeed, according to Equation 1, when a sequence is encoded with high QP, more weight is assigned to rate which leads the encoder to select larger blocks. The early termination methods included in VTM are more likely to stop the partitioning process earlier for high QP. This tends to reduce complexity further. This observation also explain why the encoding time per CTU is lower for class E. These sequences have fixed unified background which leads the encoder to select larger blocks.

To conclude, this analysis shows that the complexity of video encoding is related to the frame rate, resolution, and QP value. However, our results also bring out that encoding complexity is content dependent. Next, we analyze the complexity impact at coding tool level.

IV. COMPLEXITY REDUCTION IN VTM INTRA ENCODER

As described in Section II-B, complexity reduction techniques commonly reduce the number of tested configurations while trying to limit the degradation of the RD-cost. In the VTM encoder, an exhaustive RDO search that leads to the minimal RD-cost is performed at three nested levels: 1) CTU partitioning; 2) intra mode prediction; and 3) MTS process. The next section determines the complexity reduction opportunities at these three levels.

A. Determination of the complexity reduction opportunities

A theoretical upper limit for complexity reduction is obtained when the encoder is able to predict perfectly the best configuration and thus only this configuration is processed. Therefore, for a given level, the complexity of the search process is reduced to the minimal complexity consumption for exactly the same encoding performance. This encoding complexity sets the theoretical complexity reduction opportunities of the according level.

Let the *optimal configurations* of CTU partitioning, intra mode prediction, and MTS process be the ones with the minimum RD-cost. These configurations are determined with two-pass encoding. The first pass is unconstrained, i.e. an exhaustive RDO search is processed, and the optimal configurations of the three levels are extracted. The second pass uses the extracted optimal configurations to force the RDO
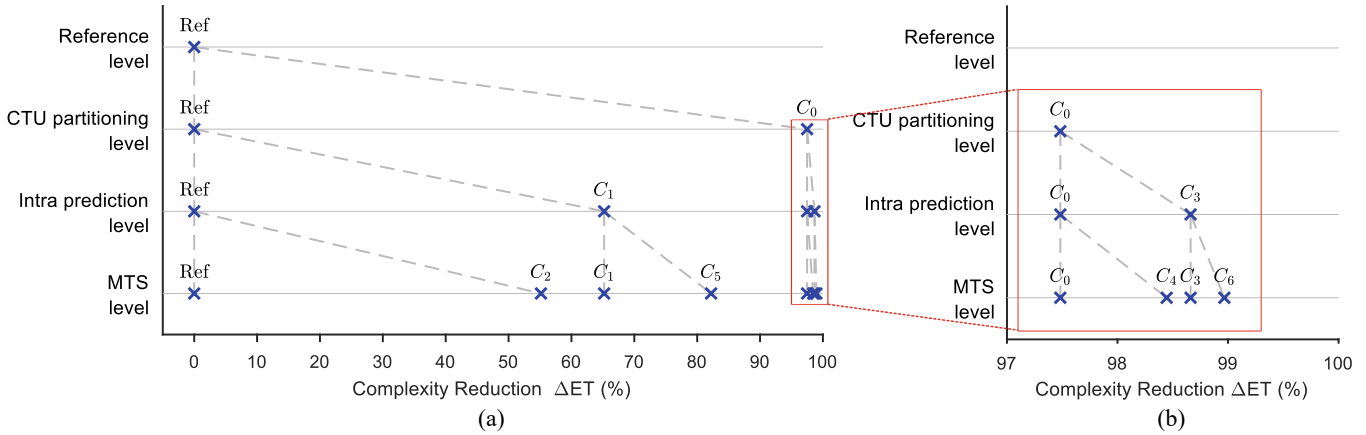
Fig. 5. Maximum complexity reduction (ΔET) of the analyzed configurations over the default configuration of VTM3.0. (a) Whole range from 0% to 100%. (b) Zoomed range from 97% to 100%.

process to encode only these configurations and thus remove unnecessary complexity. The encoding results of these two passes are identical in terms of bit rate and distortion, only the complexity differs. Finally, the complexity reduction opportunities are defined by the difference of encoding complexity Δ Encoding Time (ΔET) between the first and second pass averaged by the four QPs. ΔET is defined as

$$\Delta ET = \frac{1}{4} \sum_{QP_i \in \{22,27,32,37\}} \frac{T_R(QP_i) - T_C(QP_i)}{T_R(QP_i)}, \quad (2)$$

where T_R is the reference encoding time of the anchor VTM3.0 encoder and T_C the time of the VTM when forced to encode the optimal configurations only. The experimental setup is described in Section III-A.

B. Analysis of complexity reduction opportunities

Table I presents seven analyzed configurations (C_0-C_6) which feature different combinations of the three levels defined in Section IV-A. For each configuration, the exhaustive search is either enabled (E) or disabled (D). When the exhaustive search is disabled at one level, only the optimal configurations of this level are performed. Following Equation 2, Ref represents the reference encoding complexity (exhaustive search) that is used as reference encoding time T_R to compute the complexity reduction ΔET. C_0, C_1 and C_2 represent the complexity reduction opportunities of the CTU partitioning, intra mode prediction and MTS levels, respectively. The other configurations are combinations of the previous ones.

Fig. 5(a) shows complexity reduction of these seven configurations organized at the three nested levels. The complexity reduction results are computed as in Equation 2 and averaged for the 22 sequences. The reference (Ref) is used as an anchor. As explained in the previous section, all configurations produce identical bit-rate and distortion, i.e, BD-BRs difference between the reference and the other configurations equals 0. Fig. 5(b) focuses on the results between 97% and 100% of

TABLE I
ANALYZED CONFIGURATIONS WITH MAXIMUM AVAILABLE COMPLEXITY
REDUCTION

Level		Configuration							
	Ref	C_0	C_1	C_2	C_3	C_4	C_5	C_6
CTU partitioning	E	D	E	E	D	D	E	D
Intra mode prediction	E	E	D	E	D	E	D	D
MTS process	E	E	E	D	E	D	D	D
ΔET (%)	0	97.48	65.20	55.16	98.66	98.45	82.21	98.97

D: Disabled, E: Enabled.

Fig. 5(a) as the points are too close to be distinguished. The average standard deviation of the results presented in Fig. 5(a) and (b) is equal to 2.11% (with a maximum of 4.89%) which confirms that the average results across QPs and classes are representative. To interpret the following results, it is important to notice that the experiments are performed under the VVC reference software VTM3.0, which is not implemented to be a practical real-time encoder.

As shown by the C_0 point on Fig. 5(a), restraining the encoder to process only the optimal configuration for the CTU partitioning level offers the best complexity reduction opportunity among the three levels, up to 97.5%. In other words, being able to perfectly predict the CTU partitioning without testing the unnecessary splits would reduce the encoding time down to 2.5% of the reference. This result is due to the multitude of partitioning possibilities introduced in VTM, including QT, BT, and TT, as described in Section II-A. The RDO process, including the search of intra mode and transform, is applied for each CU. Reducing the number of intra modes tested by the RDO process for all CUs can reduce complexity by up to 65.2% over the reference with 67 modes, as presented by the C_1 point on Fig. 5(a). Finally, the MTS process enables a complexity reduction opportunity of 55.2% when the optimal horizontal and vertical transforms can be predicted for all TUs. The MTS in VTM tests several transform for each TU that significantly increases the encoding time.

Since the aforementioned three levels are nested but independent, it is possible to process only the optimal configurations of multiple levels simultaneously. Reducing the complexity of multiple levels is interesting as it was previously done on several works in HEVC [10], [11]. As shown by the C_5 point on Fig. 5(a), addressing the intra mode prediction and MTS together offers a complexity reduction opportunity of up to 82.2%. This result is in line with the individual complexity reduction opportunities of the intra prediction and MTS level $(65.2\% + (100\% - 65.2\%) \times 55.2\% = 84.4\% \approx 82.2\%)$. As shown in Fig. 5(b) with the C_3, C_4 and C_6 points, when intra prediction and MTS levels are combined with the CTU partitioning level, the complexity reduction opportunities are not much higher than that of the CTU partitioning level alone (less than 2% of difference). Indeed, if the CTU partitioning is predicted perfectly, the complexity used by the RDO process to select the intra prediction mode and transform is very low as the RDO process is only done on the selected CUs.

Considering the current version of VTM3.0, we can conclude that the complexity reduction issue can be more efficiently addressed by optimizing the CTU partitioning process rather than the intra mode prediction or MTS process. Furthermore, knowing the theoretical upper limits of complexity reduction at different levels of VVC may help to better evaluate current and future optimization techniques. To illustrate that, a method cutting 50% of the MTS complexity has almost reached its maximum whereas the same result on the CTU partitioning is only half the maximum.

V. CONCLUSION

Complementary to state-of-the-art techniques, this paper analyzed complexity reduction opportunities of VVC intra encoder by using a hierarchical approach, from encoding parameters to encoding tool level. This study demonstrated that the relative complexity of VTM encoding is proportional to video resolution and QP. Furthermore, the main contribution of this paper was to bring out the theoretical upper limits for various complexity reduction techniques. The CTU partitioning is shown to have potential for complexity reduction of up to 97% whereas the respective percentages for intra mode prediction and MTS are 65% and 55%.

The results of this study will pave the way for optimizations in practical VVC encoder implementations. Our future work will extend this study further by evaluating the complexity reduction opportunities in VVC inter coding.

REFERENCES

[1] CISCO, "Cisco Visual Networking Index : Forecast and Trends, 2017-2022," Tech. Rep., 2019.

[2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[3] B. Bross, J. Chen, and S. Liu, "Versatile Video Coding (Draft 5)," *JVET-N1001*, Mar. 2019.

[4] F. Bossen, X. Li, and K. Suehring, "AHG report: Test model software development (AHG3)," *JVET-M0003*, Jan. 2019.

[5] J. Ma, A. Wieckowski, V. George, T. Hinz, J. Brandenburg, S. De-Luxán-Hernández, H. Kichhoffer, R. Skupin, H. Schwarz, D. Marpe, T. Schierl, and T. Wiegand, "Quadtree plus binary tree with shifting (including software)," *JVET-J0035*, Apr. 2018.

[6] F. Racapé, G. Rath, F. Urban, L. Zhao, S. Liu, X. Zhao, x. Li, A. Filippov, v. Rufitskiy, and J. Chen, "CE3-related: Wide-angle intra prediction for non-square blocks," *JVET-K0500*, Jul. 2018.

[7] W. Chien, J. Boyce, R. Chernyak, K. Francois, R. Hashimoto, Y. He, Y. Huang, and S. Liu, "JVET AHG report: Tool reporting procedure," *JVET-L0013*, Oct. 2018.

[8] A. Mercat, F. Arrestier, W. Hamidouche, M. Pelcat, and D. Menard, "Energy reduction opportunities in an HEVC real-time encoder," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 1158–1162.

[9] Z. Liu, X. Yu, Y. Gao, S. Chen, X. Ji, and D. Wang, "CU Partition Mode Decision for HEVC Hardwired Intra Encoder Using Convolution Neural Network," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5088–5103, Nov. 2016.

[10] G. Correa, P. A. Assuncao, L. V. Agostini, and L. A. d. S. Cruz, "Fast HEVC Encoding Decisions Using Data Mining," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 4, pp. 660–673, Apr. 2015.

[11] B. Huang, Z. Chen, Q. Cai, M. Zheng, and D. Wu, "Rate-Distortion-Complexity Optimized Coding Mode Decision for HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2019.

[12] Z. Wang, S. Wang, J. Zhang, S. Wang, and S. Ma, "Effective Quadtree Plus Binary Tree Block Partition Decision for Future Video Coding," in *2017 Data Compression Conference (DCC)*, Apr. 2017, pp. 23–32.

[13] ——, "Probabilistic Decision Based Block Partitioning for Future Video Coding," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1475–1486, Mar. 2018.

[14] Z. Jin, P. An, C. Yang, and L. Shen, "Fast QTBT Partition Algorithm for Intra Frame Coding through Convolutional Neural Network," *IEEE Access*, pp. 1–1, 2018.

[15] H. Zhang and Z. Ma, "Fast Intra Mode Decision for High Efficiency Video Coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 4, pp. 660–668, Apr. 2014.

[16] K. CHEN, X. ZENG, and Y. FAN, "CNN Oriented Fast CU Partition Decision and PU Mode Decision for HEVC Intra Encoding," in *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Oct. 2018, pp. 1–3.

[17] S. Ryu and J. Kang, "Machine Learning-Based Fast Angular Prediction Mode Decision Technique in Video Coding," *IEEE Transactions on Image Processing*, vol. 27, no. 11, pp. 5525–5538, Nov. 2018.

[18] H. Yang, L. Shen, X. Dong, Q. Ding, P. An, and G. Jiang, "Low Complexity CTU Partition Structure Decision and Fast Intra Mode Decision for Versatile Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2019.

[19] F. Bossen, J. Boyce, K. Suehring, X. Li, and V. Seregin, "JVET common test conditions and software reference configurations for SDR video," *JVET-M1010*, Jan. 2019.