Learning-Based Complexity Reduction Scheme for VVC Intra-Frame Prediction

Mário Saldanha[1], Gustavo Sanchez[2], César Marcon[3], Luciano Agostini[1]
[1]Federal University of Pelotas – Graduate Program in Computing (PPGC) – Pelotas, Brazil
[2]IF Farroupilha – Alegrete, Brazil
[3]Pontificial Catholic University of Rio Grande do Sul – Porto Alegre, Brazil
{mrdfsaldanha, agostini}@inf.ufpel.edu.br, gustavo.sanchez@iffarroupilha.edu.br, cesar.marcon@pucrs.br

2021 International Conference on Visual Communications and Image Processing (VCIP) | 978-1-7281-8551-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/VCIP53242.2021.9675394

***Abstract*—This paper presents a learning-based complexity reduction scheme for Versatile Video Coding (VVC) intra-frame prediction. VVC introduces several novel coding tools to improve the coding efficiency of the intra-frame prediction at the cost of a high computational effort. Thus, we developed an efficient complexity reduction scheme composed of three solutions based on machine learning and statistical analysis to reduce the number of intra prediction modes evaluated in the costly Rate-Distortion Optimization (RDO) process. Experimental results demonstrated that the proposed solution provides 18.32% encoding timesaving with a negligible impact on the coding efficiency.**

Keywords—Complexity Reduction, Intra Prediction, Machine Learning, Versatile Video Coding.

I. INTRODUCTION

Versatile Video Coding (VVC) [1] is the most recent international video coding standard developed by the Joint Video Experts Team (JVET). VVC allows higher coding efficiency than High Efficiency Video Coding (HEVC) [2] and provides increased versatility in a wide range of applications. VVC introduces several novel tools and improvements, such as Quadtree with nested Multi-type Tree (QTMT) [3], new intra- and inter-frame prediction tools, Multiple Transform Selection (MTS) [4], and Low-Frequency Non-Separable Transform (LFNST) [5].

VVC intra-frame prediction brings several new coding tools and substantial modifications in those inherited from HEVC, improving the coding efficiency. VVC supports 65 angular prediction modes instead of the 33 supported by HEVC. The intra-frame prediction can be performed for square and rectangular block shapes using the Wide-Angle Intra Prediction (WAIP) [6] approach. New tools, like Multiple Reference Line (MRL) [7], Intra Subpartition (ISP) [8], and Matrix-based Intra Prediction (MIP) [9], also contribute to VVC achieving higher coding efficiency in intra-frame prediction at an expressive additional computational cost. Bossen et al. [10] showed that VVC Test Model (VTM) [11] provides a 25% higher coding efficiency than HEVC Test Model (HM) at the cost of 26 times more encoding time, considering the All-Intra (AI) encoder configuration.

This fact motivated the development of novel solutions [12]-[19] to reduce the encoding complexity while keeping the coding efficiency. Most of these works [12]-[17] proposed fast Coding Unit (CU) decisions to reduce the encoding complexity of the QTMT partitioning structure, whereas the works [18] and [19] proposed fast intra prediction mode decisions. Even though these works presented impressive results in terms of timesaving and coding efficiency, most of them focus only on the QTMT structure, and the intra mode decisions are limited to the angular prediction modes. These works focused only on the angular prediction modes since they targeted the first VTM versions that did not include several new standardized tools. In the current VTM version, these fast intra prediction mode decisions tend to provide low timesaving due to the new intra-frame prediction tools. Therefore, there is room to develop solutions for reducing the coding complexity of the new intra-frame prediction modes, contributing to the real-time VVC processing.

In this paper, we propose a novel learning-based complexity reduction scheme for VVC intra-frame prediction composed of three solutions: (i) a fast Planar/DC decision based on Decision Tree (DT) classifier, (ii) a fast MIP decision based on DT classifier, and (iii) a fast ISP decision based on the block variance. These three solutions result in an efficient mode decision scheme to discard intra prediction modes from the Rate-Distortion (RD) list, avoiding the evaluation of prediction modes that are unlikely to be chosen as the best ones in the costly Rate-Distortion Optimization (RDO) process. The proposed scheme can reduce the encoding time significantly with a negligible impact on the coding efficiency.

II. VVC INTRA-FRAME PREDICTION

VVC intra-frame prediction introduces several innovations, especially for luminance blocks. Regarding the conventional intra-frame prediction approach used in HEVC, VVC extends the HEVC angular prediction modes from 33 to 65 angular modes. The Planar and DC modes remain with the same approach used in HEVC. VVC enables the intra-frame prediction for rectangular block shapes using the WAIP approach, where prediction modes with angles beyond 45 degrees are evaluated according to the block size ratio [6]. MRL [7] allows the use of two farther reference lines for the VVC intra-frame prediction.

MIP [9] is an alternative approach to the conventional angular intra prediction modes, performing the intra-frame prediction through matrix multiplication and sample interpolation. A set of matrices were defined according to the block size by offline training through neural networks, and each matrix represents a prediction mode.

ISP [8] explores the correlation among the block samples to improve the conventional intra-frame prediction. ISP divides horizontally or vertically the current block into two or four subpartitions, which are sequentially encoded using the same prediction mode. The reconstructed samples of each encoded subpartition are used as reference samples for the next subpartition.

Fig. 1 presents the encoding flow of VVC intra-frame prediction for luminance blocks. The encoder evaluates several encoding modes to minimize the Rate-Distortion cost (RD-cost). Then, the encoding process selects the prediction mode that reaches the lowest RD-cost.

The VTM intra-frame prediction employs Rough Mode Decision (RMD) and Most Probable Modes (MPM) [20] approaches to build a list of promising candidates referred to as RD-list. RMD estimates the encoding cost of each candidate mode based on the Sum of Absolute Transformed

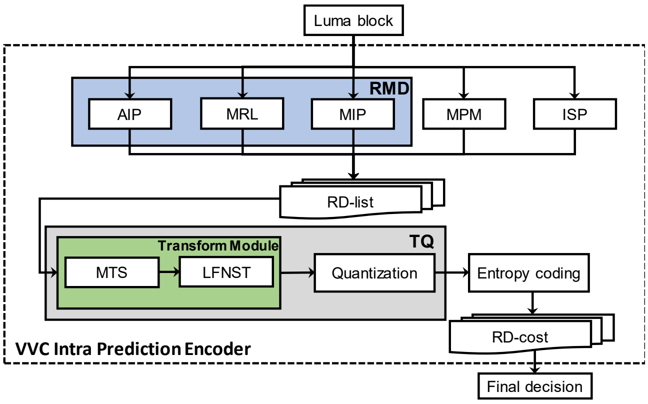


Fig. 1. Encoding flow of the VVC intra-frame prediction for luminance.

Differences (SATD), then a few modes with the lowest costs are ordered and inserted in the RD-list. After, MPM gets the default modes (the most frequently used ones) and the modes in the left and above neighbor blocks and inserts at most two additional modes into the RD-list. The RMD process is applied to the Angular Intra Prediction (AIP), MRL, and MIP tools. Then, the MPMs are inserted at the end of the RD-list. For simplifying, we call AIP the conventional intra-frame prediction composed of Planar, DC, 65 angular modes, and wide-angle prediction modes.

ISP does not use the RMD process since this tool requires the real reconstructed samples used as a reference to get the next subpartition prediction. Hence, VTM adopts some strategies to derive the most promising prediction modes [21], inserting the ISP modes in the RD-list and evaluating in the RDO process only after computing the RD-costs of AIP, MRL, MIP, and MPM tools. Then, all modes in the RD-list are processed by the full RDO process, applying residual and entropy coding. The residual coding encompasses Transform and Quantization (TQ) steps. The transform module includes MTS [4] and LFNST [5], which are tools for primary and secondary transforms, respectively. Subsequently, the RD-costs are obtained and the prediction mode with the lowest RD-cost is selected as the best mode for the current CU. RDO is a complex process since it includes the residual coding, entropy coding, and the reconstruction of reference samples, then efficient solutions to reduce the number of candidate modes in the RD-list that will be evaluated in RDO are required to decrease the encoding time of VVC intra-frame prediction.

A. *Motivational Analysis*

To develop efficient solutions for reducing the encoding time while maintaining the coding efficiency, we have made two analyses showing the mode selection percentage of the VVC intra coding tools. These analyses were performed in the VTM 10.0 following the Common Test Conditions (CTC) [22] specified by JVET for AI encoder configuration.

Fig. 2 presents the occurrence of VVC intra coding tools, considering the four Quantization Parameter (QP) values and the average results for all video sequences defined in CTC. This analysis evaluates the mode selection distribution between the conventional prediction tools (AIP+MRL) and the new intra coding tools (MIP and ISP). This evaluation combines the occurrence of AIP and MRL (AIP+MRL) since they use the same prediction approach, differing only on the used the reference samples. One can notice that AIP+MRL is more used than MIP and ISP for all cases evaluated. On average, AIP+MRL is selected 68% of the time, followed by MIP and ISP that are selected 24% and 8% of the time, respectively. The QP variation presents slight variations in this distribution.

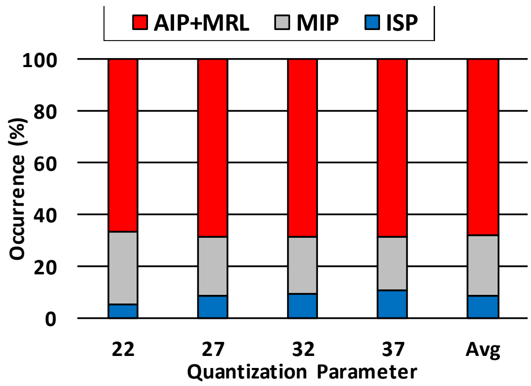


Fig. 2. Mode selection distribution for intra coding tools considering four QPs.

Fig. 3 presents the occurrence of the prediction modes in the conventional intra-frame prediction approach, considering the average results for all QPs and video sequences defined in CTC. Planar and DC modes are represented by 0 and 1, respectively, and angular modes are in the range 2 to 66. Planar and DC modes are frequently used to encode smooth texture regions, whereas angular modes encode more complex texture patterns. One can notice that Planar and DC modes are the most selected prediction modes, occurring about 43% of the time. Among the angular modes, the most selected one (mode 50) occurs less than 5% of the time.

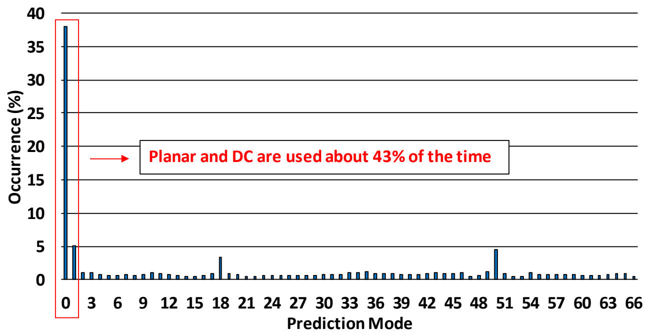


Fig. 3. Mode selection distribution for conventional intra-frame prediction approach.

This result allows us to conclude that, for most cases, the encoding selects the conventional intra-frame prediction approach. Besides, when the conventional intra-frame prediction is selected, Planar and DC modes are frequently used in the encoding process. Therefore, an intelligent complexity reduction scheme to decide when to avoid the evaluation of angular, MIP, and/or ISP modes in the RDO process can provide high encoding timesaving with negligible impact on the coding efficiency.

III. Learning-Based Complexity Reduction Scheme

The inclusion of new tools in the VVC intra-frame prediction significantly increased the number of candidate modes from the RD-list to be evaluated in the time-consuming RDO process, requiring solutions that expressively reduce time while maintaining coding efficiency. This section presents the learning-based complexity reduction scheme, which can discard intra prediction modes from the RD-list before the RDO process. This scheme is composed of three solutions: (i) a fast Planar/DC decision based on DT classifiers, (ii) a fast MIP decision based on DT classifiers, and (iii) a fast ISP decision based on the block variance.

A. Methodology

We employed data mining to discover strong correlations between the encoding context and its attributes for defining DT classifiers that determine when discarding some prediction modes from the RD-list. For this purpose, we implemented further VTM functions and collected statistical information from the encoding process to train the DT classifiers offline using the REPTree algorithm in the *Waikato Environment for Knowledge Analysis* (WEKA) [23]. The datasets were balanced according to the number of instances for each frame, block size, QP value, and output class.

We used eight video sequences in the training process with resolutions ranging from 416×240 to 3840×2160 pixels: TrafficFlow [24], BuildingHall2 [24], ParkScene [24], Kimono1 [24], Vidyo1 [24], Netflix_DrivingPOV [25], pedestrian_area (downsampled to 832×480) [26], and Flowervase [24]. These video sequences include a wide range of video characteristics for rendering examples of mode decisions in the training process. All experiments were executed using the VTM 10.0 following the CTC parameters for AI configuration and considering the QP values 22, 27, 32, and 37.

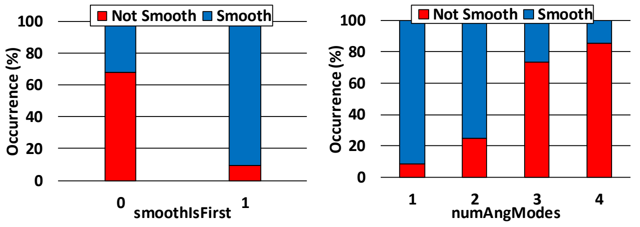
B. Fast Planar/DC Decision based on DT Classifier

Since the conventional intra-frame prediction selects Planar and DC for many cases, which are frequently used to encode smooth texture regions, our first solution identifies when the smooth modes (Planar and DC) tend to be selected as the best prediction modes and, in this case, our solution removes the angular modes from the RD-list, avoiding their evaluations in the RDO. For this purpose, we collected a large amount of data from the encoding process and defined a DT classifier for deciding when the RDO evaluation of the angular modes can be avoided since these modes are unlikely to be chosen as the best ones. Table I displays the attributes used in the DT classifier and the corresponding descriptions.

Table I – Attributes used in the Planar/DC DT classifier.

Attribute	Description
QP	The current QP value
area	The current block area
dcInRdList	Notify if DC is on RD-list
posPlanar	The position of the Planar in the RD-list
posDC	The position of the DC in the RD-list
dcIsMPM	Notify if DC is an MPM
smoothIsFirst	Notify if Planar or DC is on the first RD-list position
mipIsFirst	Notify if MIP mode is on the first RD-list position
numAngModes	Number of angular modes in the RD-list
numMipModes	Number of MIP modes in the RD-list
numModesList	Total number of modes in the RD-list
dcSATDCost	SATD-cost of DC
planarSATDCost	SATD-cost of Planar
firstSATDNoSmooth	SATD-cost of the first angular or MIP in the RD-list
ratioSATDSmooth	SATD-cost of Planar divided by SATD-cost of DC
ratioSATD	Minimum SATD-cost between Planar and DC divided by firstSATDNoSmooth

As examples, Fig. 4(a) and (b) show the correlation of *smoothIsFirst* and *numAngModes* attributes with the smooth mode decision, respectively. *Not Smooth* indicates when an angular mode is selected as the best mode, and *Smooth* specifies when Planar or DC is selected as the best mode. When *smoothIsFirst* is equal to 1, a smooth mode is selected about 91% of the time. Also, the smaller the value of *numAngModes*, the higher is the probability of a smooth mode be selected, demonstrating a high correlation of these attributes with the smooth mode decision.



(a) (b)

Fig. 4. Correlation between *smoothIsFirst* and *numAngModes* attributes with the smooth mode decision.

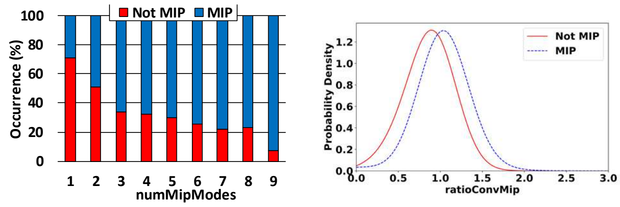
C. Fast MIP Decision based on DT Classifier

The analysis presented in Section II.A shows that MIP modes are selected only about 24% of the time. Therefore, our second solution identifies when MIP modes can be discarded from the RD-list since they are unlikely to be selected in the RDO. This solution also collects a large amount of data from the encoding process to define a DT classifier for deciding when MIP modes can be removed from the RD-list, avoiding the RDO evaluation. Table II displays the attributes used in this DT classifier and the corresponding descriptions.

Table II. Attributes used in the MIP DT classifier.

Attribute	Description
QP	The current QP value
width	The current block width
height	The current block height
area	The current block area
blockRatio	Block width divided by block height
QTMTD	The current QTMT depth level
posPlanar	The position of the Planar in the RD-list
posFirstMip	The position of the first MIP in the RD-list
numMipModes	Number of MIP modes in the RD-list
numModesList	Total number of modes in the RD-list
planarSATDCost	SATD-cost of Planar
firstSATDMip	SATD-cost of the first MIP in the RD-list
ratioConvMip	SATD-cost of the first conventional mode divided by the first MIP in the RD-list
numNeighMip	Number of neighboring blocks encoded with MIP

Fig. 5(a) and (b) exemplify the correlation between *numMipModes* and *ratioConvMip* with the MIP decision, respectively. *Not MIP* and *MIP* indicate when a non-MIP or MIP mode is selected, respectively. The lower the value of *numMipModes*, the higher the probability of a non-MIP to be selected. Considering *ratioConvMip*, low values or values close to 1 tend to select a non-MIP mode while values higher than 1 tend to choose a MIP mode. Both attributes demonstrate a high correlation with the MIP mode decision.

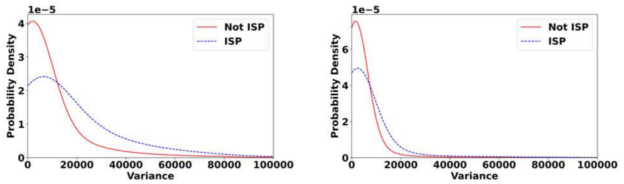


(a) (b)

Fig. 5. Correlation between two attributes and the MIP decision.

D. Fast ISP Decision based on the Block Variance

ISP is frequently used to encode more complex texture regions where the conventional intra-frame prediction approach cannot provide high coding efficiency. However, the conventional prediction can provide better rate-distortion results for simpler texture regions, allowing to avoid ISP evaluation. Based on this fact, our third solution analyzes the block texture complexity to decide when to remove ISP modes from the RD-list. Fig. 6 shows the probability density functions for the block variance when the current block is predicted with ISP or not, considering Vidyo1 and Kimono1 video sequences encoded with QP 32.



(a) Vidyo1 (b) Kimono1

Fig. 6. Probability density function for ISP mode selection considering the block variance of different video sequences.

One can notice that the block variance can provide a good indication of when applying the ISP prediction and a threshold decision can be defined to remove ISP modes from the RD-list. These figures also demonstrate that a static threshold definition may deal with inaccurate decisions for some cases since it tends to vary according to the video content and QP value. Therefore, our solution uses the block variance to decide when removing the ISP modes, but the threshold value is computed online during the encoding of the first frame of the video sequence. Our solution stores the variance values of all blocks of the first frame that did not select ISP as the best mode and computes the average variance value. This variance value is used in the subsequent frames as a threshold to remove the ISP evaluations in the RDO. Besides, this threshold value can be periodically adjusted according to the application requirements, considering the number of frames, change of video content or, target timesaving.

IV. Experimental Results

This section presents the results of the proposed scheme for VVC intra coding. We performed all the experiments in VTM 10.0 following the CTC specified by JVET [22] for AI encoder configuration. CTC encompasses six classes of video sequences with resolutions ranging from 416×240 up to 3840×2160 pixels and four QP values (22, 27, 32, and 37).

It is important to highlight that the training process for DT classifiers did not use JVET CTC video sequences; consequently, this evaluation considered different video sequences from the ones used in the training step, allowing a robust evaluation of the proposed scheme. Besides, VTM implements some native speedup heuristics for the intra-frame prediction process, such as a fast RMD search of the 65 angular modes, the evaluation of the MRL for only six MPMs, and fast decisions for ISP based on the previous prediction mode evaluations as described in [21]. In our experiments, all these speedup techniques are enabled, allowing a fairer comparison with the current implementation of the VTM encoder.

Both Planar/DC and MIP DTs were designed with tree depth eight and evaluated using 10-fold cross-validation, achieving accuracy results of 85.37% and 78.83%, respectively. Table III presents the Encoding Time Saving (ETS) and Bjontegaard Delta Bitrate (BDBR) [27] results of Planar/DC and MIP DTs, fast ISP, and overall proposed scheme. Planar/DC and MIP DTs reach 10.47% of ETS with a small increase of 0.29% in BDBR. Fast ISP reduces the coding complexity by 8.32% increasing the BDBR by 0.31%. Combining all solutions, the proposed scheme can provide an ETS of 18.32% with a negligible BDBR increase of 0.60%. Besides, the proposed scheme presents a small standard deviation (σ), demonstrating stable results for different video content and resolutions. The highest ETS (20.84%) was obtained in the MarketPlace video sequence with a BDBR increase of only 0.22%. The highest BDBR impact (1.11%) was attained with BasketballDrill video sequence but still providing 16.05% of ETS. The ETS and BDBR results range from 12.63% to 20.84% and from 0.18% to 1.11%, respectively.

Table III. Proposed scheme results for CTC under AI configuration.

Video Sequence	Planar/DC and MIP DTs		Fast ISP		Overall	
	BDBR (%)	ETS (%)	BDBR (%)	ETS (%)	BDBR (%)	ETS (%)
Tango2	0.50	11.90	0.08	7.92	0.56	18.89
FoodMarket4	0.29	8.60	0.07	6.58	0.36	15.29
Campfire	0.26	11.82	0.03	9.26	0.29	19.63
CatRobot	0.33	9.94	0.24	7.01	0.58	17.62
DaylightRoad2	0.39	12.80	0.24	9.23	0.64	19.71
ParkRunning3	0.13	8.15	0.07	5.24	0.18	12.63
MarketPlace	0.11	12.47	0.08	10.67	0.22	20.84
RitualDance	0.19	9.92	0.24	6.84	0.44	19.79
Cactus	0.26	12.05	0.29	5.94	0.55	19.36
BasketballDrive	0.42	10.20	0.43	7.95	0.84	19.36
BQTerrace	0.36	10.29	0.37	10.22	0.72	20.07
BasketballDrill	0.41	6.25	0.69	7.99	1.11	16.05
BQMall	0.33	10.31	0.66	9.43	0.94	19.40
PartyScene	0.19	11.26	0.34	11.18	0.54	20.28
RaceHorsesC	0.15	11.47	0.26	11.90	0.41	18.17
BasketballPass	0.36	9.98	0.41	7.46	0.72	16.08
BQSquare	0.32	10.52	0.39	8.25	0.69	19.24
BlowingBubbles	0.23	9.96	0.34	7.09	0.57	16.55
RaceHorses	0.15	11.02	0.28	11.38	0.43	19.86
FourPeople	0.28	11.48	0.46	6.65	0.79	18.21
Johnny	0.45	9.61	0.48	6.92	0.85	17.30
KristenAndSara	0.38	10.32	0.43	7.83	0.83	18.75
Average	**0.29**	**10.47**	**0.31**	**8.32**	**0.60**	**18.32**
σ	**0.11**	**1.50**	**0.18**	**1.85**	**0.24**	**1.98**

To the best of our knowledge, this is the first complexity reduction solution for VVC intra coding considering all the novel intra-frame prediction tools; consequently, it is difficult to perform a fair comparison with related works. The works [18] and [19] used an old version of VTM (2.0) and reached 25.51% of ETS with a 0.54% BDBR increase and 30.59% of ETS with a 0.86% BDBR increase, respectively. Since our scheme targeted the VTM with all standardized tools, having a more complex process to build the RD-list, one can conclude that our solution can provide high coding time savings with a minimum impact on the coding efficiency.

V. Conclusions

This paper presented a learning-based complexity reduction scheme to reduce the VVC intra coding time. The proposed scheme comprises three fast decisions capable of avoiding the evaluation of angular, MIP, and/or ISP prediction modes in the costly RDO process. Two DT classifiers can decide when avoiding the evaluation of angular and MIP modes, whereas the ISP evaluation is avoided by measuring the texture variance of the current block. The proposed solution can provide 18.32% of timesaving with negligible impact on the coding efficiency of only 0.60%, contributing to enable VVC real-time encoding.

Acknowledgment

The authors thank FAPERGS, CNPq and CAPES (Finance Code 001) Brazilian research support agencies that financed this investigation.

REFERENCES

[1] ITU-T and ISO/IEC, "Versatile Video Coding," ITU-T Rec. H.266 and ISO/IEC 23090-3, 2020.

[2] G. J. Sullivan, J. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits Syst. Video Technol. (TCSVT)*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012, doi: 10.1109/TCSVT.2012.2221191.

[3] Y. Huang et al., "A VVC Proposal with Quaternary Tree Plus Binary-Ternary Tree Coding Block Structure and Advanced Coding Techniques," *IEEE Trans. Circuits Syst. Video Technol. (TCSVT)*, vol. 30, no. 5, pp. 1311-1325, May 2020, doi: 10.1109/TCSVT.2019.2945048.

[4] X. Zhao, J. Chen, M. Karczewicz, L. Zhang, X. Li and W. Chien, "Enhanced Multiple Transform for Video Coding," *Data Compression Conf. (DCC)*, Snowbird, UT, 2016, pp. 73-82, doi: 10.1109/DCC.2016.9.

[5] M. Koo, M. Salehifar, J. Lim and S. Kim, "Low Frequency Non-Separable Transform (LFNST)," *IEEE Picture Coding Symp. (PCS)*, Ningbo, China, 2019, pp. 1-5, doi: 10.1109/PCS48520.2019.8954507.

[6] L. Zhao et al., "Wide Angular Intra Prediction for Versatile Video Coding," *2019 Data Compression Conf. (DCC)*, Snowbird, UT, USA, 2019, pp. 53-62, doi: 10.1109/DCC.2019.00013.

[7] B. Bross et al. "CE3: Multiple reference line intra prediction (Test 1.1.1, 1.1.2, 1.1.3 and 1.1.4)," JVET 12th Meeting, JVET-L0283, Macao, Oct. 2018.

[8] S. De-Luxán-Hernández et al., "An Intra Subpartition Coding Mode for VVC," *IEEE Int. Conf. Image Process. (ICIP)*, Taipei, Taiwan, 2019, pp. 1203-1207, doi: 10.1109/ICIP.2019.8803777.

[9] M. Schäfer et al., "An Affine-Linear Intra Prediction with Complexity Constraints," IEEE Int. Conf. Image Process. (ICIP), Taipei, Taiwan, 2019, pp. 1089-1093, doi: 10.1109/ICIP.2019.8803724.

[10] F. Bossen, X. Li, K. Suehring, K. Sharman, V. Seregin, A. Tourapis, "AHG report: Test model software development (AHG3)," JVET 22nd Meeting, JVET-V0003, teleconference, Apr. 2021.

[11] VVC Test Model (VTM). Available at: https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM, access in June 2021.

[12] M. Lei, F. Luo, X. Zhang, S. Wang and S. Ma, "Look-Ahead Prediction Based Coding Unit Size Pruning for VVC Intra Coding," *IEEE Int. Conf. Image Process. (ICIP)*, 2019, pp. 4120-4124, doi: 10.1109/ICIP.2019.8803421.

[13] J. Cui, T. Zhang, C. Gu, X. Zhang and S. Ma, "Gradient-Based Early Termination of CU Partition in VVC Intra Coding," *Data Compression Conf. (DCC)*, 2020, pp. 103-112, doi: 10.1109/DCC47342.2020.00018.

[14] F. Chen, Y. Ren, Z. Peng, G. Jian and X. Cui, "A fast CU size decision algorithm for VVC intra prediction based on support vector machine". *Multimed Tools Appl* 79, 27923–27939, 2020.

[15] A. Tissier, W. Hamidouche, J. Vanne, F. Galpin and D. Menard, "CNN Oriented Complexity Reduction Of VVC Intra Encoder," *IEEE Int. Conf. Image Process. (ICIP)*, 2020, pp. 3139-3143, doi: 10.1109/ICIP40778.2020.9190797.

[16] Y. Li, G. Yang, Y. Song, H. Zhang, X. Ding and D. Zhang, "Early Intra CU Size Decision for Versatile Video Coding Based on a Tunable Decision Model," *IEEE Trans. Broadcasting (TBC)*, doi: 10.1109/TBC.2021.3073556.

[17] T. Li, M. Xu, R. Tang, Y. Chen and Q. Xing, "DeepQTMT: A Deep Learning Approach for Fast QTMT-based CU Partition of Intra-mode VVC," *IEEE Transactions on Image Processing (TIP)*, doi: 10.1109/TIP.2021.3083447.

[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 Trans. Circuits Syst. Video Technol. (TCSVT)*, vol. 30, no. 6, pp. 1668-1682, June 2020, doi: 10.1109/TCSVT.2019.2904198.

[19] Y. Chen, L. Yu, H. Wang, T. Li and S. Wang, "A novel fast intra mode decision for versatile video coding," *J. Vis. Commun. Image Representation*, v. 71, p. 102849, 2020, doi: 10.1016/j.jvcir.2020.102849.

[20] L. Zhao, L. Zhang, S. Ma and D. Zhao, "Fast mode decision algorithm for intra prediction in HEVC," *IEEE Vis. Commun. Image Process. (VCIP)*, Tainan, Taiwan, 2011, pp. 1-4, doi: 10.1109/VCIP.2011.6115979.

[21] S. De-Luxán-Hernández et al., " Design of the intra subpartition mode in VVC and its optimized encoder search in VTM," *Applications of Digital Image Processing XLIII*, v. 11510, p. 115100Y, 2020, doi: 10.1117/12.2567171.

[22] F. Bossen, J. Boyce, X. Suehring, X. Li, V. Seregin, "JVET common test conditions and software reference configurations for SDR video," JVET 14th Meeting, JVET-N1010, Mar. 2019.

[23] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. Witten, "The WEKA Data Mining Software: An Update," *ACM SIGKDD Explorations Newsletter*, v. 11, n. 1, pp. 10-18, Jan. 2009.

[24] K. Sharman and K. Suehring, "Common test conditions", JCT-VC 26th meeting, JCTVC-Z1100, Jan. 2017.

[25] T. Daede, A. Norkin and I. Brailovskiy, "Video codec testing and quality measurement," draft-ietf-netvc-testing-08 (work in progress), p. 23, 2019.

[26] Xiph.org Video Test Media [derf's collection]. Xiph.org. Acessed on: June 2021. [Online]. Available: https://media.xiph.org/video/derf/.

[27] G. Bjontegaard, "Calculation of Average PSNR Differences between RDCurves," VCEG Meeting, pp. 1-5, 2001.