Fast Transform Decision Scheme for VVC Intra-Frame Prediction Using Decision Trees

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

Abstract—This paper presents a fast transform decision scheme for Versatile Video Coding (VVC) intra-frame prediction using decision trees. VVC introduces several novel coding tools to improve the coding efficiency of the intra-frame prediction at the cost of a high computational effort, including a new transform coding process using Multiple Transform Selection (MTS) for primary transform and Low-Frequency Non-Separable Transform (LFNST) for secondary transform. We developed an efficient complexity reduction scheme composed of two solutions based on decision tree classifiers to avoid the MTS and LFNST evaluations in the costly Rate-Distortion Optimization (RDO) process. Experimental results showed that the proposed scheme provides 11% of encoding timesaving with a negligible impact on the coding efficiency.

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

I. Introduction

Versatile Video Coding (VVC) [1] is the most recent international video coding standard, which was collaboratively developed by the Joint Video Experts Team (JVET). VVC was designed to provide an impressive bitrate reduction over High Efficiency Video Coding (HEVC) [2] standard with high video quality and high versatility for handling different video content and applications.

VVC follows the same concept of block-based hybrid video coding approach as its predecessor, HEVC. Each frame of a video sequence is split into blocks, and these blocks are processed by intra- and inter-frame prediction, transform, quantization, entropy coding, and in-loop filters. VVC improves the compression performance by inserting several novel tools and enhancements for each of these steps. These improvements include bigger block sizes and a large number of available block sizes, Quadtree with nested Multi-type Tree (QTMT) partitioning [3], Multiple Reference Line (MRL) [4][5] intra-frame prediction, Matrix-based Intra Prediction (MIP) [6], Multiple Transform Selection (MTS) [7], Low-Frequency Non-Separable Transform (LFNST) [8], Luma Mapping with Chroma Scaling (LMCS) [9], among others.

All these techniques contribute to significantly improve the VVC coding efficiency. Bossen et al. [10] reported that VVC Test Model (VTM) [11] provides 25% higher coding efficiency than HEVC Test Model (HM) for All-Intra (AI) encoder configuration. However, while these tools enhance the coding efficiency, the encoder computational burden rises expressively. The VTM intra coding complexity increased by 26 times [10] when compared to HM for AI configuration.

This high coding complexity boosted novel and efficient complexity reduction solutions for VVC intra coding [12]-[20]. Most of these works [12]-[18] designed fast Coding Unit (CU) decisions to reduce the encoding complexity of the QTMT partitioning structure, whereas the works [18] and [19] developed fast intra-frame prediction mode decisions, and the work [20] proposed a fast transform selection decision for the intra-frame prediction. Although these works presented very good timesaving and coding efficiency results, there is still much room for VVC encoder complexity improvement since most of these works focused on QTMT structure and the works [18], [19], and [20] targeted the first VTM versions that did not include some VVC standardized tools.

This paper proposes a novel fast transform decision scheme for VVC intra-frame prediction of luminance samples using Decision Trees (DTs). This scheme comprises two solutions: (i) a fast MTS decision based on the DT classifier and (ii) a fast LFNST decision based on the DT classifier. These two solutions result in an efficient scheme to avoid the evaluation of primary and secondary transforms 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 with a negligible impact on the coding efficiency.

II. Transform Coding in VVC Intra-frame Prediction for Luminance Samples

VVC enhances the transform coding by including MTS [7] and LFNST [8], which are tools for primary and secondary transform modules, respectively. MTS is available for intra- and inter-frame predicted blocks, whereas LFNST is available only for intra-frame predicted blocks. Fig. 1 presents the transform coding process for the VVC intra-frame prediction of luminance samples, which is the focus of this work.

The intra-prediction encoder locally evaluates several combinations of prediction modes to create a list of promising prediction modes, avoiding the evaluation of all modes through the costly RDO process. This list is referred as Rate-Distortion list (RD-list). After the RD-list definition, the transforms are applied and VVC allows a combination of different transforms intending to minimize the RD-cost. Besides the Discrete Cosine Transform II (DCT-II), used in HEVC, VVC also allow the use of Discrete Sine Transform

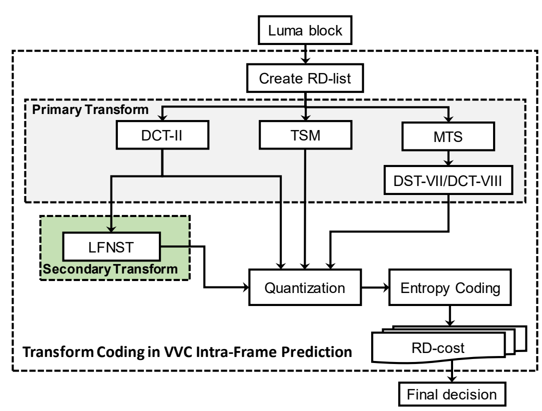

Fig. 1. Transform coding flow in the VVC intra-frame prediction of luminance samples.

VII (DST-VII) and DCT-VIII, increasing the contribution of this module to the global coding efficiency, but also increasing a lot the computational effort required to process the transforms, if compared with the HEVC. VVC also defines the use of Transform Skip Mode (TSM), as in HEVC. TSM is available for 32×32 or smaller blocks and, in this case, the prediction residues are directly sent to the quantization step, avoiding the use of transforms.

VVC transform coding has three main processing paths regarding the primary transform application, as presented in Fig. 1: (i) the first one using DCT-II for horizontal and vertical directions; (ii) the second one using TSM, and (ii) the third one using the MTS, where DST-VII and DCT-VIII are used. The paths (i) and (ii) are similar to the HEVC transforms. The use of MTS in path (iii) allows a combination of DST-VII and DCT-VIII in horizontal and vertical directions, then four combinations are evaluated: (i) DST-VII and DST-VII, (ii) DST-VII and DCT-VIII, (iii) DCT-VIII and DST-VII, and (iv) DCT-VIII and DCT-VIII [21]. The VTM implementation sequentially process these steps, where the DCT-II is firstly processed, followed by the TSM and the MTS calculation is the last processing step.

The primary transform block sizes are different where DCT-II has sizes ranging from 4×4 to 64×64, and DST-VII and DCT-VIII have sizes ranging from 4×4 to 32×32. In all cases, rectangular shapes are allowed.

An exception for this processing flow of the primary transforms is the Intra Subpartition [22] tool. In this case, combinations between DCT-II and DST-VII are allowed and the decision about which transform will be used in horizontal or vertical direction is done considering the width and height of the processed sub-partition [21].

When the DCT-II is used, then the application of the LFNST secondary transform is allowed, as presented in Fig. 1. This transform further decorrelates the low-frequency primary transforms coefficients. LFNST may be applied for transform blocks ranging from 4×4 to 64×64, including rectangular shapes and contains two transform sets (LFNST 1 and LFNST 2) with four non-separable transform matrices for each set [8]. The transform matrix evaluated for each set is defined based on the used intra-frame prediction mode [21]. The VTM process the evaluation of DCT-II without LFNST, DCT-II with LFNST 1, followed by DCT-II with LFNST 2. When LFNST is not applied, the DCT-II results are sent directly to quantization.

The encoder evaluates many possibilities through the complex RDO process to find the combination of intra prediction mode, primary, and secondary transform with the lowest RD-cost. Although the VTM encoder implements some fast decisions for this process, as described in [21], efficient solutions able to identify when MTS and/or LFNST tools are unlikely to be chosen as the best ones are desired to reduce RDO evaluations and decrease the encoding time of the VVC intra-frame prediction.

III. Motivational Analysis

To better evaluate the transform coding process in the VVC intra-frame prediction and to develop efficient solutions to reduce the encoding time while maintaining the coding efficiency, we have performed two analyses about the transform selection percentage regarding MTS and LFNST coding tools. These analyses were performed in the VTM 10.0 following the Common Test Conditions (CTC) [23] defined by JVET for AI encoder configuration, considering the average results for all video sequences and Quantization Parameter (QP) values.

Fig. 2 presents the occurrence of the primary transform when encoded with MTS (using DST-VII/DCT-VIII) or with the other options (DCT-II or TSM), regarding all 17 block sizes available in the VVC intra-frame prediction and regarding the global average. 64×64 blocks are never encoded with MTS since this tool is not available for this block size. On average, MTS is selected 35% of the time, indicating that DCT-II/TSM are selected for most cases. The distribution changes a little bit according with the block sizes, but the with small standard deviation, as one can notice from Fig. 2.

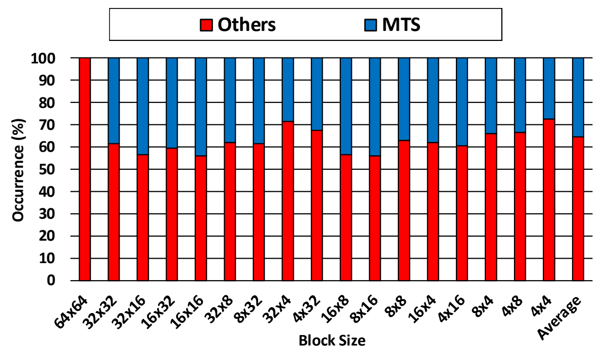

Fig. 2. MTS occurrence for VVC intra-frame prediction.

Fig. 3 presents the occurrence of the secondary transform for each block size when the DCT-II is applied as the primary transform. In this case, “With LFNST” is when LFNST 1 or LFNST 2 are applied and “Without LFNST” is when the DCT-II results are sent directly to quantization. On average, LFNST is used in 49% of the blocks, demonstrating that more than 50% of the cases are encoded without secondary transform. In this experiment, different block sizes show a higher variation in the results if compared with Fig. 2. The lowest and the highest percentage of LFNST use happen with 64×64 (33%) and 16×16 (66%) blocks sizes, respectively.

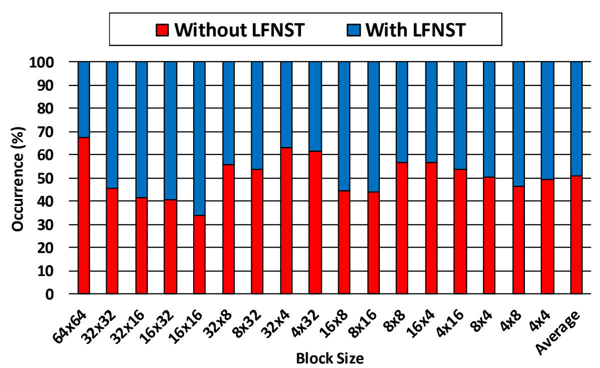

Fig. 3. LFNST occurrence for VVC intra-frame prediction.

The presented analyses allow us to conclude that, in most cases, the blocks are encoded without the use of MTS or LFNST coding tools. Consequently, the evaluations of MTS and LFNST in the costly RDO process can be avoided for several cases. Therefore, an efficient complexity reduction scheme able to decide when to avoid the evaluation of MTS and/or LFNST in the RDO can provide interesting timesaving with negligible impact on the coding efficiency.

IV. Proposed Solution

This section presents the proposed fast transform decision scheme based on DT classifiers, which can avoid some of the evaluations of MTS and LFNST coding tools.

A. Methodology

We used data mining to discover strong correlations between the encoding context and its attributes for defining DT classifiers that determine when removing the evaluation of MTS and/or LFNST tools. We implemented further VTM functions and collected statistical information from the encoding process to train the DT classifiers using the REPTree algorithm in the Waikato Environment for Knowledge Analysis (WEKA) [24]. We employed DT classifiers since it presents more hardware-friendly characteristics compared to the other machine learning classifiers.

We used eight non-CTC video sequences in the training process with resolutions ranging from 416×240 to 3840×2160 pixels: TrafficFlow [25], BuildingHall2 [25], ParkScene [25], Kimono1 [25], Vidyo1 [25], Netflix_DrivingPOV [26], pedestrian_area (downsampled to 832×480) [27], and Flowervase [25]. These video sequences include a wide range of video characteristics for rendering several examples of transform decisions in the training process. The datasets were balanced according to the number of instances for each frame, block size, QP value, and output class. All experiments were performed with the VTM 10.0 following the CTC parameters for AI configuration, considering the QPs 22, 27, 32, and 37.

B. Fast MTS Decision based on DT Classifier

Since MTS is selected only about 35% of the time, and it is evaluated later than DCT-II and TSM (Others) in the VTM implementation, the coding information obtained prior to MTS evaluation can be used to detect when MTS evaluation is unnecessary. Based on the encoding context and the current coding information, our first solution identifies when the MTS evaluation can be skipped from the RDO process. We collected a large amount of data from the encoding process and defined a DT classifier for deciding when the MTS evaluation can be skipped. Table I demonstrates the attributes used in the DT classifier, the corresponding descriptions and Information Gain (IG) [28].

Table I. Attributes used in the MTS DT classifier.

Attribute	Description	IG
QP	The current QP value	0.016
area	The current block area	0.086
width	The current block width	0.116
BTD	The current binary tree depth	0.111
mipFlag	Notify if MIP was selected	0.009
ispMode	Identify the ISP mode	0.017
currCost	RD-cost	0.054
currDistortion	Total distortion	0.073
currFracBits	Number of encoded bits	0.090
currIntraMode	Intra prediction mode	0.068
noIspCost	RD-cost of the best non-ISP mode	0.117
numNonZeroCoeffs	Number of non-zero coefficients	0.026
absSumCoeffs	Absolute sum of the coefficients	0.094
numNeighMTS	Number of neighboring blocks encoded with MTS	0.030

Fig. 4(a) and (b) exemplify the correlation between *numNeighMTS* and *ispMode* with the MTS decision, respectively. When the *numNeighMTS* value is zero, about 61% of the blocks are encoded without MTS. In contrast, the higher the value of *numNeighMTS*, the higher the probability of encoding a block with MTS. Considering *ispMode*, when this attribute has the value one or two, more than 62% of the blocks are encoded without MTS, whereas if the value is zero, 51% of the blocks are encoded with MTS. Both attributes demonstrate a correlation with the MTS decision.

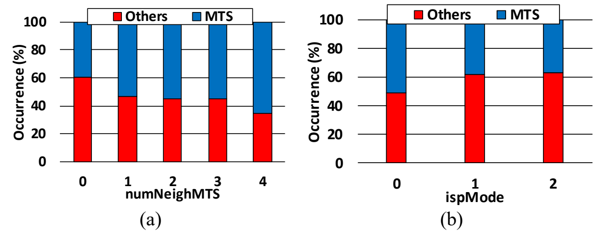

Fig. 4. Correlation between *numNeighMTS* and *ispMode* attributes with the MTS decision.

C. Fast LFNST Decision based on DT Classifier

Since from the blocks encoded with DCT-II, more than 50% are encoded without LFNST, our second solution identifies when the LFNST evaluation can be skipped. For this purpose, we collected a large amount of data from the encoding process when the DCT-II is evaluated without LFNST. Based on the collected data, we trained a DT classifier to decide when the LFNST evaluation can be avoided from the RDO process. Table II displays the attributes used in this DT classifier, the corresponding descriptions and IG.

Table II. Attributes used in the LFNST DT classifier.

Attribute	Description	IG
width	The current block width	0.019
height	The current block height	0.170
area	The current block area	0.055
blockRatio	Block width divided by block height	0.021
MTTD	The current multi-type tree depth	0.176
mipFlag	Notify if MIP was selected	0.106
ispMode	Identify the ISP mode	0.015
currCost	RD-cost	0.086
currDistortion	Total distortion	0.136
currFracBits	Number of encoded bits	0.087
currIntraMode	Intra prediction mode	0.013
numNonZeroCoeffs	Number of non-zero coefficients	0.189
absSumCoeffs	Absolute sum of the coefficients	0.046

Fig. 5(a) and (b) show the correlation between *MTTD* and *ispMode* attributes with the LFNST decision, respectively. When the *MTTD* value is zero, about 69% of the blocks are encoded without LFNST. In contrast, for higher values of *MTTD*, the percentage of blocks encoded with LFNST is about 52% (value 4). Regarding *ispMode*, this attribute presents a similar behavior when compared to the MTS DT classifier. When *ispMode* is one or two, more than 67% of the blocks are encoded without LFNST, and when this attribute has the value zero, 50% of the blocks are encoded with LFNST, demonstrating a correlation with the LFNST decision.

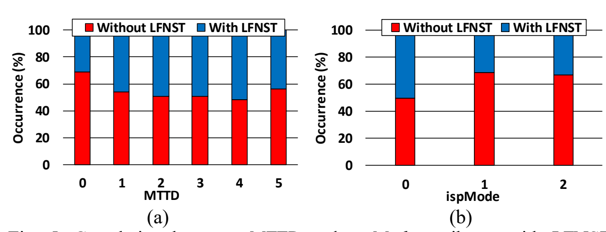

Fig. 5. Correlation between *MTTD* and *ispMode* attributes with LFNST decision.

D. Fast Transform Decision Scheme

Fig. 6 presents the flowchart of the fast transform decision scheme using DT classifiers. After creating the RD-list, the encoder evaluates DCT-II (without LFNST) and TSM mode. Subsequently, based on the result of the transform coding with the lowest RD-cost, we compute the features for the MTS and LFNST DT classifiers (Sections IV.B and IV.C) and also compute the decision of the MTS and LFNST DTs. When the MTS DT classifier decides to skip the MTS evaluation, the RDO complexity is reduced; otherwise, no simplification is performed and the MTS is evaluated. Analogously, when the LFNST DT classifier decides to skip the LFNST evaluation, the RDO process is simplified; otherwise, the LFNST coding evaluation remains without modifications.

V. EXPERIMENTAL RESULTS

This section presents the results obtained with the proposed fast transform decision scheme. This scheme was implemented inside the VTM 10.0 and evaluated following the CTC specified by JVET for AI encoder configuration [23]. CTC encompasses six classes of video sequences with resolutions ranging from 416×240 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, as previously discussed. Consequently, this evaluation considered different video sequences from the ones used in the training step, allowing an unbiased evaluation of the proposed scheme. Besides, VTM implements some native speedup heuristics for the transform coding process, such as deciding to skip MTS and LFNST evaluations according to the ISP results and terminate the MTS evaluation according to the results of a certain MTS mode, 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 MTS and LFNST DTs were designed with tree depth eight and have size of 125 and 123, respectively. The MTS and LFNST DTs were evaluated using 10-fold cross-validation, achieving accuracy results of 73.87% and 76.75%, respectively.

Table III presents the Encoding Time Saving (ETS) and Bjontegaard Delta Bitrate (BDBR) [29] results of MTS DT, LFNST DT, and the overall proposed scheme. MTS DT reaches 5% of ETS with a small increase of 0.21% in BDBR. LFNST DT reduces the coding complexity by 6.40%, increasing the BDBR by 0.23%. The proposed scheme can provide an ETS of 10.99% with a negligible BDBR increase of 0.43% when combining both solutions. Besides, the proposed scheme presents a small standard deviation (σ), demonstrating stable results for different video content and resolutions. The ETS and BDBR results range from 7.81% to 13.30% and from 0.15% to 0.70%, respectively.

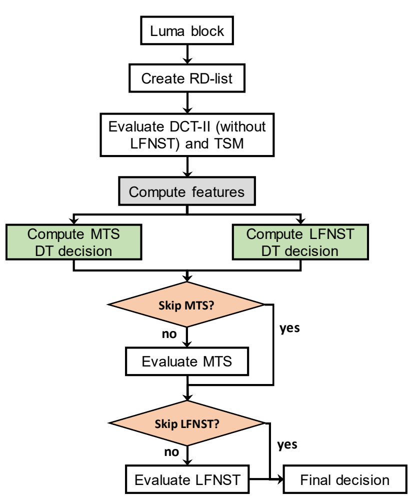

Fig. 6. Flowchart of the fast transform decision scheme.

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

Video Sequence	MTS DT		LFNST DT		Overall	
	BDBR (%)	ETS (%)	BDBR (%)	ETS (%)	BDBR (%)	ETS (%)
Tango2	0.22	6.04	0.47	6.29	0.70	11.83
FoodMarket4	0.27	2.72	0.41	5.26	0.68	8.91
Campfire	0.17	3.25	0.16	4.21	0.32	8.24
CatRobot	0.24	5.65	0.18	6.75	0.40	9.94
DaylightRoad2	0.20	6.21	0.19	3.88	0.40	11.92
ParkRunning3	0.21	2.76	0.12	6.67	0.30	7.81
MarketPlace	0.24	5.29	0.14	5.54	0.38	13.17
RitualDance	0.27	5.37	0.32	7.62	0.60	7.96
Cactus	0.23	3.89	0.19	3.90	0.40	11.52
BasketballDrive	0.27	5.45	0.27	6.85	0.49	11.40
BQTerrace	0.11	4.45	0.16	5.61	0.25	11.75
BasketballDrill	0.25	6.03	0.47	4.64	0.64	9.98
BQMall	0.25	7.13	0.17	8.45	0.41	10.44
PartyScene	0.10	5.93	0.08	9.58	0.20	12.70
RaceHorsesC	0.17	4.81	0.18	7.10	0.32	13.14
BasketballPass	0.27	5.50	0.20	6.72	0.45	9.51
BQSquare	0.09	4.39	0.07	6.68	0.15	13.30
BlowingBubbles	0.14	5.20	0.12	7.81	0.26	12.20
RaceHorses	0.17	4.91	0.17	8.64	0.34	11.74
FourPeople	0.33	4.29	0.36	7.51	0.66	11.23
Johnny	0.28	4.67	0.33	4.60	0.58	10.49
KristenAndSara	0.21	6.13	0.35	6.53	0.59	12.55
Average	**0.21**	**5.00**	**0.23**	**6.40**	**0.43**	**10.99**
σ	**0.06**	**1.14**	**0.12**	**1.57**	**0.16**	**1.70**

To the best of our knowledge, this is the first complexity reduction solution for VVC transform coding considering all the novel transform coding tools for VVC intra-frame prediction; consequently, it is difficult to perform a fair comparison with related works. The work [20] used an old version of VTM (3.0) and reached 23% of ETS with a 0.16% BDBR increase; however, this solution also reduces the number of intra prediction modes in the RD-list and does not consider the LFNST coding tool. Since our scheme targeted VTM with all standardized tools (e.g., LFNST, MIP, and ISP), having a more complex process to decide the best encoding possibility, one can conclude that our solution can provide high coding time savings with a minimum impact on the coding efficiency. Besides, our scheme also can be combined with solutions to reduce the number of intra prediction modes in the RD-list or fast CU decisions to reach even more impressive timesaving results.

VI. CONCLUSIONS

This paper presented a fast transform coding decision scheme using DT classifiers to reduce the VVC intra coding time. The proposed scheme comprises two fast decisions capable of avoiding the evaluation of MTS and/or LFNST coding tools in the costly RDO process. For this purpose, two DT classifiers were offline trained with effective features to deal with the MTS and LFNST decisions. The proposed scheme can provide 11% of timesaving with negligible impact on the coding efficiency, enabling 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. -W. Huang et al., "Block Partitioning Structure in the VVC Standard," *IEEE Trans. Circuits Syst. Video Technol. (TCSVT)*, doi: 10.1109/TCSVT.2021.3088134.

[4] 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.

[5] J. Pfaff et al., "Intra Prediction and Mode Coding in VVC," *IEEE Trans. Circuits Syst. Video Technol. (TCSVT)*, doi: 10.1109/TCSVT.2021.3072430.

[6] 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.

[7] 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.

[8] 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.

[9] T. Lu et al., "Luma Mapping with Chroma Scaling in Versatile Video Coding," *Data Compression Conf. (DCC)*, Snowbird, UT, USA, 2020, pp. 193-202, doi: 10.1109/DCC47342.2020.00027

[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] T. Fu, H. Zhang, F. Mu and H. Chen, "Two-Stage Fast Multiple Transform Selection Algorithm for VVC Intra Coding," *IEEE Int. Conf. Multimedia Expo (ICME)*, Shanghai, China, 2019, pp. 61-66, doi: 10.1109/ICME.2019.00019.

[21] X. Zhao et al., "Transform Coding in the VVC Standard," *IEEE Trans. Circuits Syst. Video Technol. (TCSVT)*, doi: 10.1109/TCSVT.2021.3087706.

[22] 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.

[23] 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.

[24] 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.

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

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

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

[28] T. M. Cover and J. A. Thomas, Elements of Information Theory, New York: John Wiley & Sons, 1991.

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