Intra Prediction in the Emerging VVC Video Coding Standard

Alexey Filippov*, Vasily Rufitskiy*, Jianle Chen+, and Elena Alshina#

Data Compression Conference 2020

Versatile Video Coding also known by its abbreviation “VVC” is an emerging International standard that is being developed by the Joint Video Experts Team (JVET) of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). At the 10th JVET meeting in April 2018 after reviewing the responses to the Call for Proposals (CfP) [1] issued in October 2017, 13 Core Experiments (CEs) were established to study proposed techniques. Among these Core Experiments, CE3 devoted to intra prediction and mode coding was one of the most long-running (it has been closed only at the 16th JVET meeting in October 2019) and active in terms of the number of tools studied. The CE3 work resulted in adopting a set of new tools into the VVC specification draft text [2] and the VVC Test Model (VTM) [3]. In this paper, we provide an overview of some video coding techniques related to the activity of CE3.

The focus of our work is mainly on such intra-prediction tools that distinguish VVC from its predecessors (e.g., AVC/H.264 [4] and HEVC/H.265 [5]). It is worth noting that a part of VVC intra prediction tools such as Cross-component linear model (CCLM) prediction [6], Matrix-based Intra Prediction (MIP) [7], Intra Subpartition (ISP) coding mode [8], Multiple Reference Line (MRL) intra prediction [9] are thoroughly covered in already published works. Hence, these methods are out of the scope of this paper. Sections below describe directional and non-directional intra prediction modes designs, respectively.

Directional intra-prediction

One of the tendencies observable for VVC as a whole is higher accuracy of the tools still used in HEVC. It relates to the new coding structure, motion vector precision, and intra prediction where the number of directional modes available for a given block is extended up to 65 against 33 HEVC directions. In addition to accuracy increase, directional intra prediction is harmonized with the new coding structure that enables blocks of both square and rectangular shapes.

Wide-angular intra prediction

One of the VVC tools heavily contributed to compression performance improvement is a flexible partitioning framework based on quad-tree (QT), binary (BT) [10] and ternary trees (TT) combined into multi-type tree (MTT) [11]. It can generate coding structure endowed with superior segmentation capabilities that provide high-precision object localization. In addition to blocks of square shape, this mechanism uses rectangular blocks that introduce asymmetry into reference sample distribution between top and left sides. If the subranges of directional modes assigned to each side cover the equal angles in the case of square blocks, the asymmetry caused by rectangular share requires these subranges to be adjusted according to the block aspect ratio (Fig. 1). More top-right prediction directions are allocated for it. Consequently, more bottom-left intra prediction modes are specified for blocks with height greater than width. Since the additional modes allocated along a longer side have an angle of greater than 45° relative to horizontal or vertical mode (subject to which of them is closer to the additional mode), this method is referred to as wide-angle intra prediction (WAIP) [12]. Fig. 2 summarizes all the directions that can appear in blocks of different orientations and aspect ratios.

It is remarkable that introduction of WAIP modes does not cause any changes in the intra mode coding. WAIP modes are correspondingly mapped to the almost opposite directions (with a minor offset by one mode) [12]. In Fig. 1, modes with indices in the range [67, 72] are correspondingly signaled as mode indices in the range [2, 7] and shown in parenthesis. Mapping of a mode index out of the range enabled for intra mode coding to the complete range presented in Fig. 2 is invoked immediately before decoding process for intra prediction. It is worth noting that restricting the range of directional modes available in a block by collinear modes belonging to a counter-diagonal is a design intention. For rectangular blocks, one of these collinear modes is a WAIP integer-slope mode out of the following mode list: -14, -12, -10, -6, 72, 76, 78, 80 [13].

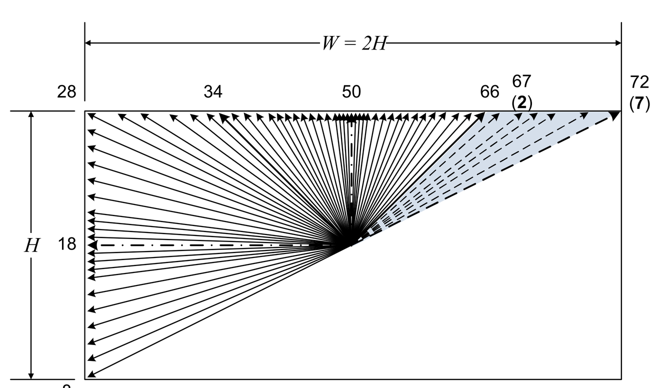


Figure 1: Set of intra directions available in a horizontally oriented block with aspect ratio of 2:1

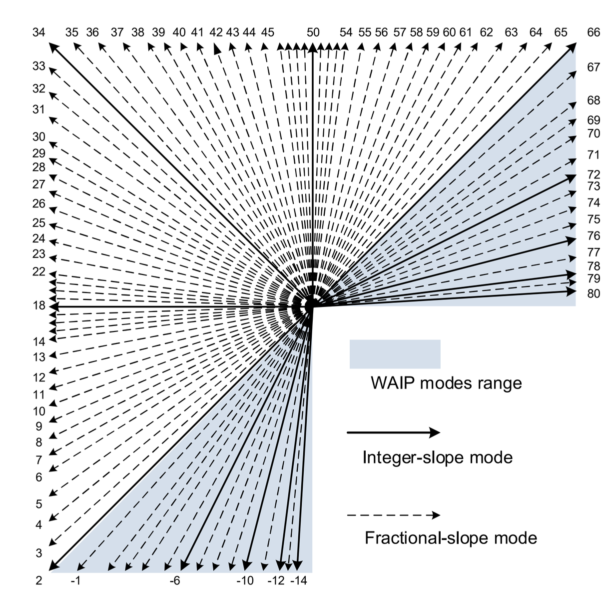
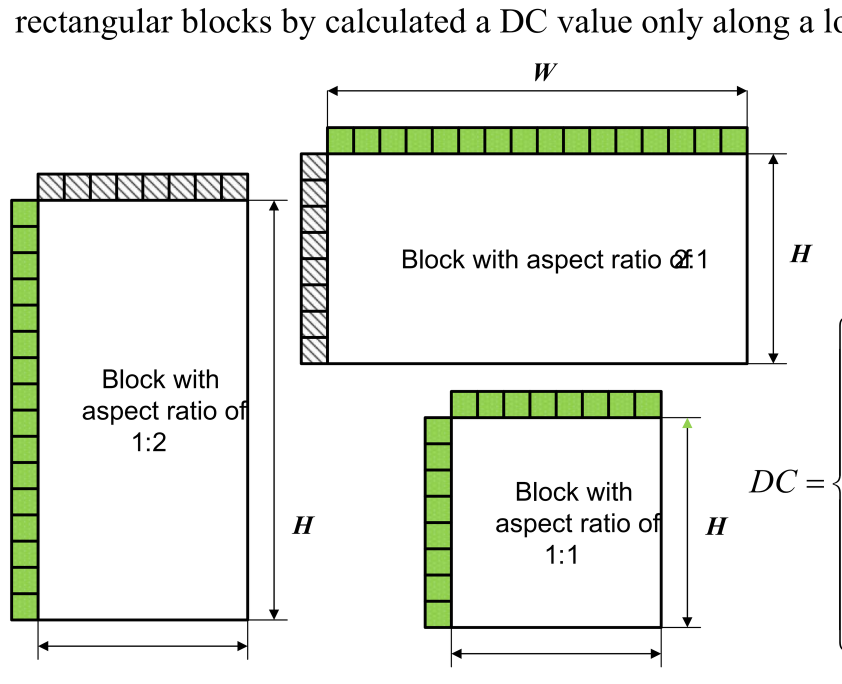


Figure 2: Extension of intra prediction directions up to 93 in total

Non-directional intra prediction

DC mode

Design of an intra prediction mode that is considered to be the least complicated has to be kept hardware-friendly for both square and rectangular blocks. In particular, division operation trivially implemented by right-shift operation only for square shape can be an issue in the case of rectangular shape. To skip this computationally complex operation, a couple of methods were proposed. Fig. 7 illustrates the technique adopted into VVC [19]. This solution, on the one hand, provides the design consistent with HEVC and, on the other hand, naturally addresses the problem occurred in the case of rectangular blocks by calculated a DC value only along a longer side



$$DC=\begin{cases}\left(\left(\sum_{j=0}^{W-1}p_{j,-1}\right)+(W>>1)\right)>>w, \text{ if } W>H\\ \left(\left(\sum_{i=0}^{H-1}p_{-1,i}\right)+(H>>1)\right)>>h, \text{ if } W<H\\ \left(\left(\sum_{j=0}^{W-1}p_{j,-1}\right)+\left(\sum_{i=0}^{H-1}p_{-1,i}\right)+H\right)>>(h+1), \text{ if } W==H\end{cases}$$

Figure 7: Calculation of a DC value for square and non-square blocks

Reference sample filtering

Similar to HEVC, intra prediction in VVC has 2 filtering mechanisms applied to reference samples, namely reference sample smoothing and interpolation filtering (see Fig. 3) [14]. Reference sample smoothing applied only to integer-slope modes in luma blocks modifies reference samples by using the finite impulse response filter [1, 2, 1]/4 that does not require multiplication operations and allows keeping the design of integer-slope modes simple as compared to fractional-slope ones, which use an interpolation filter. It is worth noting that such a design enables multiplication-free implementations of “lazy” encoders if among directional modes, only horizontal, vertical and integer-slope modes are checked. More detailed explanations on handling reference samples could be found in [15].

The interpolation filter set includes 3 filters [16]. Two of them are applied to luma blocks:

- 4-tap DCT-based interpolation filter (DCT-IF) that is identical to chroma DCT-IF used for motion compensation in both HEVC and VVC [15];
- 4-tap smoothing interpolation filter (SIF) that is obtained by convolving linear filter with [1, 2, 1]/4 to be consistent with reference sample smoothing [16].

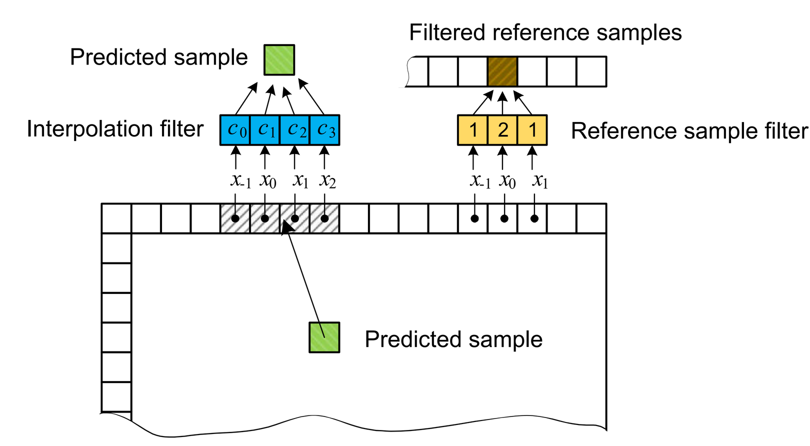


Figure 3: Reference sample smoothing and interpolation filtering

These two filters are switched based on the rules known as mode-dependent intra smoothing (MDIS) that are similar to the ones used in HEVC. The frequency responses of both filters are shown in Fig. 4. Chroma blocks are interpolated using conventional linear filter. None of these three filters changes reference samples. Note that the VVC design intention was to avoid two sequentially invoking reference sample smoothing and interpolation filtering that could result in a latency issue.

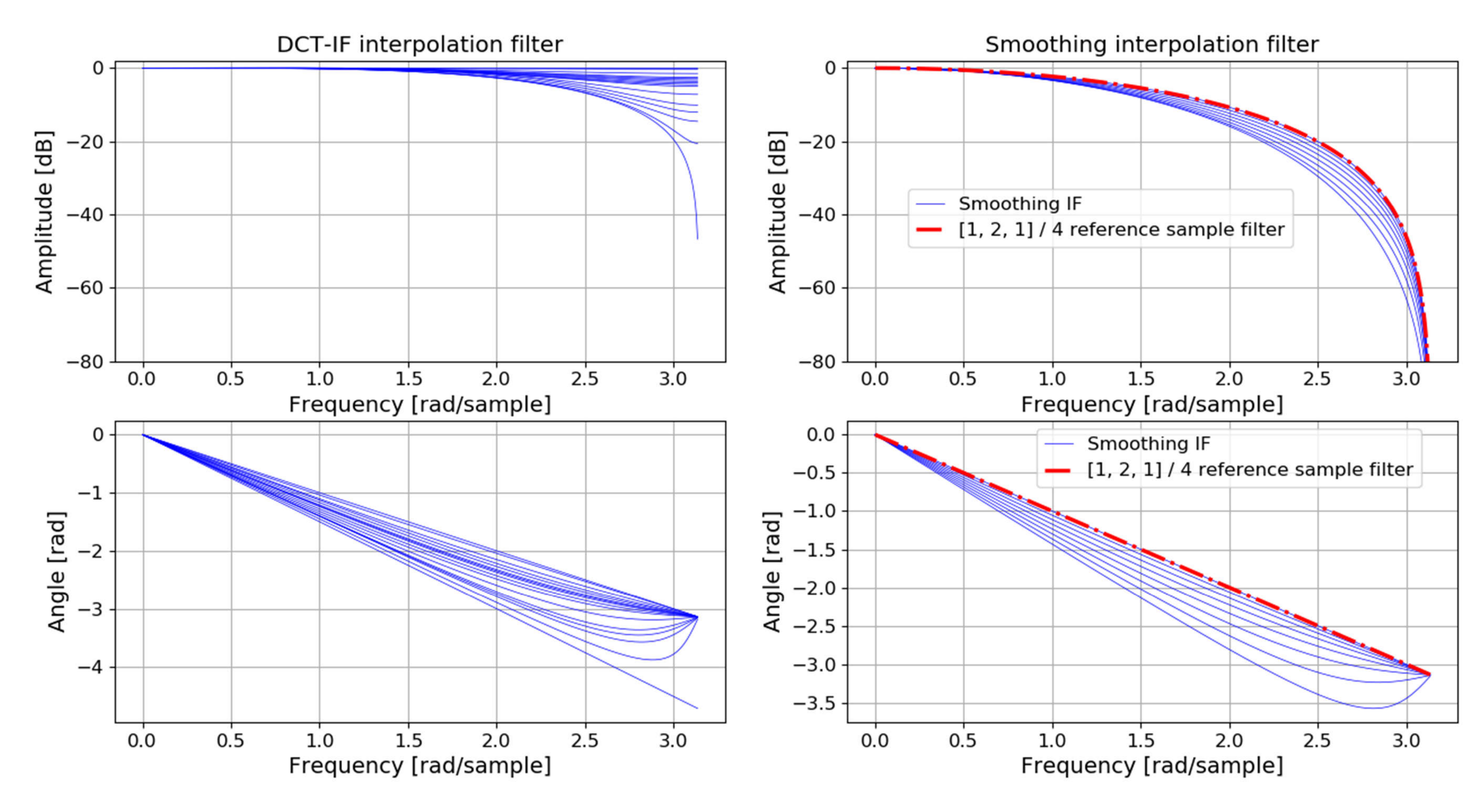


Figure 4: Frequency responses of interpolation and reference filters used in VVC

PDPC

Yet, as observed for intra prediction in AVC/H.264 and HEVC/H.265, both directional and non-directional modes may introduce discontinuities along block boundaries [14]. To remedy this problem, post-prediction filtering was applied to DC, horizontal and vertical modes [17]. In VVC, the number of cases that use post-prediction filtering was extended in two aspects. First, whereas only one column and row of intra-predictors are filtered in HEVC, propagation depth is increased and becomes dependent on block size (the larger block size, the deeper propagation). Second, this filter is applied to a wider subset of intra prediction modes (e.g., to modes -14..10, and 58..80) [17[18]. Thus, the post-prediction filter known as position-dependent prediction combination (PDPC) is a technique that suppresses discontinuities near a left and above boundaries within the predicted block as shown in Fig. 6. The following formula is used to update the predicted sample $p(x,y)$:

$$p(x,y) = \mathrm{Clip1}\left((\mathrm{ref}_L(x,y)\cdot wL(x) + \mathrm{ref}_T(x,y)\cdot wT(y) + (64 - wL(x) - wT(y))\cdot p(x,y) + 32) >> 6\right).$$

However, some PDPC steps are mode-specific. In the case of horizontal or vertical modes, PDPC propagates a weighted value of gradient calculated as a difference between a top-left reference sample $p(-1,-1)$ and $p(x,-1)$ or $p(-1,y)$ for horizontal or vertical modes, respectively. Thus, the equation to calculate the updated sample values for horizontal or vertical modes is as follows:

$$p(x,y) = \mathrm{Clip1}((\mathrm{R}(-1,y) - \mathrm{R}(-1,-1))\cdot wL[x] + (\mathrm{R}(x,-1) - \mathrm{R}(-1,-1))\cdot wT[y] + 64\cdot p(x,y) + 32) >> 6$$

where R(x,y) denotes reference samples, $p(x,y)$ denotes predicted samples, Clip1() is a clipping function that prevents exceeding the sample range by an output value and is actually required only for these two modes. For other modes, the updated sample is a weighted sum of predicted sample and a reference sample.

* Huawei Technologies Co., Ltd., {alexey.filippov,vasily.rufitskiy}@huawei.com

+ Futurewei Technologies, jianle.chen@futurewei.com

Huawei Technologies Düsseldorf GmbH, {elena.alshina}@huawei.com

Experimental results

In Table 1, we present the results obtained by VTM6.0 reference software for All-Intra configuration as compared with HM-16.20 in accordance with the JVET Common Test Conditions (CTC) [20]. Although not only intra prediction tools contribute to coding efficiency improvement, which overall value is appr. 24% for luma component, this table demonstrates the superior compression performance of VVC for intra coding configuration. To analyze the impact of each intra prediction tools studied in this paper, a separate test for each evaluated tool.

Table **1**. Objective performance of VTM6.0 over HM16-20.

Sequence		BD-rate			Time ratio		Sequence		BD-rate			Time ratio	
		Y	U	V	EncT	DecT			Y	U	V	EncT	DecT
A1	Tango2	-27.00%	-43.40%	-38.40%	1313%	186%	C	BasketballDrill	-31.20%	-33.80%	-37.80%	3601%	210%
	FoodMarket4	-27.00%	-24.10%	-29.40%	1340%	184%		BQMall	-21.40%	-22.30%	-26.60%	3948%	205%
	Campfire	-30.40%	-32.90%	-33.20%	2494%	191%		PartyScene	-15.80%	-16.30%	-19.00%	4780%	224%
Overall, class A1		**-28.10%**	**-33.50%**	**-33.60%**	**1637%**	**187%**		RaceHorses	-18.90%	-6.80%	-12.60%	4112%	205%
A2	CatRobot1	-27.60%	-28.80%	-26.30%	2200%	194%	**Overall, class C**		**-21.80%**	**-19.80%**	**-24.00%**	**4089%**	**211%**
	DaylightRoad2	-25.80%	-33.70%	-12.90%	2498%	191%	E	FourPeople	-24.40%	-22.80%	-25.40%	2854%	184%
	ParkRunning3	-30.20%	0.30%	0.00%	3337%	204%		Johnny	-26.20%	-21.00%	-26.80%	2139%	181%
Overall, class A2		**-27.90%**	**-20.70%**	**-13.10%**	**2637%**	**196%**		KristenAndSara	-25.00%	-20.10%	-24.70%	2210%	186%
B	MarketPlace	-19.90%	-14.10%	-23.10%	2960%	206%	**Overall, class E**		**-25.20%**	**-21.30%**	**-25.60%**	**2380%**	**184%**
	RitualDance	-22.30%	-17.90%	-29.90%	2218%	195%	F	BasketballDrillText	-33.20%	-35.40%	-38.50%	6256%	208%
	Cactus	-22.00%	-14.00%	-18.20%	3425%	192%		ArenaOfValor	-27.10%	-28.00%	-24.60%	6147%	226%
	BasketballDrive	-23.30%	-29.50%	-30.10%	2751%	207%		SlideEditing	-57.50%	-53.40%	-58.10%	4407%	168%
	BQTerrace	-18.10%	-24.00%	-35.10%	3025%	194%		SlideShow	-37.90%	-41.00%	-46.10%	2132%	178%
Overall, class B		**-21.10%**	**-19.90%**	**-27.30%**	**2847%**	**199%**	**Overall, class F**		**-38.90%**	**-39.40%**	**-41.80%**	**4360%**	**193%**
Overall, CTC:		**-24.20%**	**-22.50%**	**-25.00%**	**2696%**	**196%**							

The indiviudal tools, including WAIP, 4-tap interpolation filtering and PDPC are tested by using VTM6.0 reference software, and the results are shown in Table. 2

Table 2. Overall CTC tool-off results for WAIP, 4-tap interpolation filtering and PDPC

Tool:	**WAIP**			**4-tap IF**			**PDPC**		
BD-rate reduction, Y, U and V:	-0.32%	-0.37%	-0.42%	-0.45%	-0.54%	-0.57%	-0.97%	-0.14%	-0.01%
Enc/dec time	101% / 98%			102% / 103%			109% / 107%		

References

[1] A. Segall, V. Baroncini, J. Boyce, J. Chen and T. Suzuki, “Joint Call for Proposals on Video Compression with Capability beyond HEVC,” document JVET-H1002, 8th JVET meeting, Macao, China, October 2017.

[2] B. Bross, J. Chen, S. Liu and Y.-K. Wang, “Versatile Video Coding (Draft 7),” document JVET-P2001, 16th JVET meeting: Geneva, CH, 1–11 Oct. 2019.

[3] J. Chen, Y. Yan and S.-H. Kim “Algorithm description for Versatile Video Coding and Test Model 7 (VTM 7),” document JVET-P2002, 16th JVET meeting: Geneva, CH, 1–11 Oct. 2019.

[4] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560–576, July 2003.

[5] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[6] K. Zhang, J. Chen, L. Zhang, X. Li and M. Karczewicz, “Enhanced Cross-Component Linear Model for Chroma Intra-Prediction in Video Coding,” IEEE Transactions on Image Processing, vol. 27, issue: 8, pp. 3983 - 3997, 2018.

[7] M. Schäfer, B. Stallenberger, J. Pfaff, P. Helle, H. Schwarz, D. Marpe, T. Wiegand, "An Affine-Linear Intra Prediction With Complexity Constraints," IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1089-1093, September 2019

[8] S. De-Luxán-Hernández, V. George, J. Ma, T. Nguyen, H. Schwarz, D. Marpe and T. Wiegand, “An Intra Subpartition Coding Mode for VVC,” IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1203-1207, September 2019.

[9] Y.-J. Chang, H.-J. Jhu, H.-Y. Jian, L. Zhao, X. Zhao, X. Li, S. Liu, B. Bross, P. Keydel, H. Schwarz, D. Marpe and T. Wiegand, “Intra prediction using multiple reference lines for the versatile video coding standard,” Proceedings Volume 11137, Applications of Digital Image Processing XLII, SPIE Optical Engineering + Applications, San Diego, California, United States, September 2019.

[10] J. An, Y.-W. Chen, K. Zhang, H. Huang, Y.-W. Huang, S. Lei, “Block partitioning structure for next generation video coding,” MPEG document m37524 and ITU-T SG16 Document COM16–C966, October 2015.

[11] X. Li, H.-C. Chuang, J. Chen, M. Karczewicz, L. Zhang, X. Zhao, A. Said, "Multi-Type-Tree," document JVET-D0117, 4th JVET meeting, Chengdu, China, Oct 2016.

[12] F. Racapé, G. Rath, F. Urban, L. Zhao, S. Liu, X. Zhao, X. Li, A. Filippov, V. Rufitsky, J. Chen, "CE3-related: Wide-angle intra prediction for non-square blocks," document JVET-K0500, 11th JVET meeting, Ljubljana, Slovenia, July 2018.

[13] L. Zhao, X. Zhao, S. Liu, X. Li, " CE3-related: Unification of angular intra prediction for square and non-square blocks," document JVET-L0279, 12th JVET meeting, Macao, China, Oct 2018.

[14] J. Lainema, F. Bossen, W. Han, J. Min, K. Ugur, “Intra coding of the HEVC standard,” IEEE Transactions on Circuits and Systems for Video Technology, volume 22, issue 12, pp. 1792 – 1801, December 2012.

[15] A. Filippov, V. Rufitskiy, J. Chen, G. Van der Auwera, A. K. Ramasubramonian, V. Seregin, T. Hsieh, M. Karczewicz, “CE3: A combination of tests 3.1.2 and 3.1.4 for intra reference sample interpolation filter,” document JVET-L0628, in Proc. of 12th JVET meeting, Macao, China, October 2018.

[16] A. Filippov, V. Rufitskiy, “Non-CE3: Cleanup of interpolation filtering for intra prediction,” document JVET-P0599, 16th JVET meeting, Geneva, Switzerland, October 2019.

[17] G. Van der Auwera, V. Seregin, A. Said, A. K. Ramasubramonian, M. Karczewicz, "CE3: Simplified PDPC (Test 2.4.1)", document JVET-K0063, 11th JVET meeting, Ljubljana, Slovenia, Jul 2018.

[18] F. Bossen, “On general intra sample prediction,” document JVET-O0364, 15th JVET meeting, Gothenburg, Sweden, Jul 2019.

[19] A. Filippov, V. Rufitsky, J. Chen, " CE3-related: Alternative techniques for DC mode without division," document JVET-K0122, 11th JVET meeting, Ljubljana, Slovenia, Jul 2018.

[20] F. Bossen, J. Boyce, X. Li, V. Seregin, K. Sühring, “JVET common test conditions and software reference configurations for SDR video,” document JVET-N1010, 14th JVET meeting, Geneva, Switzerland, March 2019.