HAL
open science

Fast Depth and Mode Decision in Intra Prediction for Quality SHVC

Dayong Wang, Yu Sun, Ce Zhu, Weisheng Li, Frédéric Dufaux, Jiangtao Luo

To cite this version:

Dayong Wang, Yu Sun, Ce Zhu, Weisheng Li, Frédéric Dufaux, et al.. Fast Depth and Mode Decision in Intra Prediction for Quality SHVC. IEEE Transactions on Image Processing, 2020, 29 (1), pp.6136-6150. 10.1109/TIP.2020.2988167 . hal-02550207

HAL Id: hal-02550207

https://hal.science/hal-02550207v1

Submitted on 22 Apr 2020

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

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

Fast Depth and Mode Decision in Intra Prediction for Quality SHVC

Dayong Wang, Yu Sun, Ce Zhu, *Fellow, IEEE,* Weisheng Li, *Member, ACM,*
Frederic Dufaux, *Fellow, IEEE,*and Jiangtao Luo, *Senior Member, IEEE*

***Abstract*—Scalable High Efficiency Video Coding (SHVC) is the extension of High Efficiency Video Coding (HEVC). In intra prediction for quality SHVC, a Coding Unit (CU) is recursively divided into a quadtree-based structure from the largest 64×64 CU to the smallest 8×8 CU, in which 35 intra prediction modes and Inter-Layer Reference (ILR) mode are checked to determine the best possible mode. This leads to very high coding efficiency but also results in an extremely high coding complexity. To improve coding speed while maintaining coding efficiency, in this paper, we propose a new efficient algorithm for fast intra prediction for enhancement layer in SHVC. First, temporal and spatial correlations, as well as their correlation degrees, are combined in a Naive Bayes classifier to predict depth probabilities and skip depths with low likelihood. Second, for a given depth candidate, we combine ILR mode probability with Partial Zero Blocks (PZBs) based on the Sum of Squared Differences (SSD) to determine whether the ILR mode is the best one. In that case, we can skip intra prediction, which requires very high complexity. Third, initial Intra Modes (IMs) are obtained through Sobel operator, and are combined with the relationship between IMs and their corresponding Hadamard Cost (HC) values to predict candidate IMs in Rough Mode Decision (RMD). Then, an analytical criterion of early termination is developed based on the HC values of two neighboring IMs in the Rate-Distortion Optimization (RDO) process. Finally, we combine depth probabilities and the distribution of residual coefficients at the current depth to early terminate depth selection. The proposed scheme can significantly decrease the complexity of depth determination while reducing the complexity of mode decision for a depth candidate. Our experimental results demonstrate that the proposed scheme can achieve a speed up gain of more than 80% in average, while maintaining coding efficiency.**

***Index Terms*—SHVC, depth decision, ILR mode, intra prediction, early termination.**

D. Wang is with School of mathematics and computer science, Shangrao Normal University, Shangrao 334001, China, and with Chongqing Key Laboratory on Big Data for Bio Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China, and also with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: wangdayong@cqupt.edu.cn).

Y. Sun is with the Department of Computer Science, University of Central Arkansas, Conway, AR 72035, USA (e-mail: yusun@uca.edu).

C. Zhu is with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: eczhu@uestc.edu.cn).

W. Li is with Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: liws@cqupt.edu.cn).

F. Dufaux is with Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France (e-mail: frederic.dufaux@l2s.centralesupelec.fr).

J. Luo is with Electronic Information and Networking Research Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: Luojt@cqupt.edu.cn).

Manuscript received April 28, 2019; revised December 2, 2019.

I. Introduction

WITH the rapid development of information technology, videos are becoming ubiquitous nowadays. On one hand, with the popularity of mobile devices, users can easily access diverse video applications. On the other hand, the rapid development of networks allows for the effective transmission of videos. The diversity of video applications, such as digital TV broadcasting, video conferencing, wireless video streaming, and smart phone communications, has consistently increased. Even for the same video application, users may use different devices to access it, and thus may have different needs in terms of resolution and quality. In addition, heterogeneous networks, such as broadband networks and wireless networks, may have different characteristics. Even within the same network, the bandwidth may also change through time. Based on these considerations, it is obvious that video applications need to adapt to various devices and network bandwidths. Scalable Video Coding (SVC) is a natural and effective solution to these applications. As the scalable extension of H.264/AVC, SVC consists of a Base Layer (BL) and one or more Enhancement Layers (ELs). Through selecting an appropriate EL, SVC can adapt to a wide variety of device capabilities, network conditions, and client applications [1]. Since SVC needs to encode multiple layers and perform inter-layer prediction, its encoding process is very complex.

The rapidly growing demand for high-resolution video applications and services has stimulated the development of the next generation video coding standard, i.e., High Efficiency Video Coding (HEVC). Through adding more advanced features and higher-efficiency coding tools, HEVC has higher compression performances than previous video coding standards. However, its encoding coding complexity is about two to four times higher than that of H.264/AVC [2]. In order to accommodate different device capabilities, network conditions, and client applications, Scalable High Efficiency Video Coding (SHVC), the scalable extension of HEVC, has also been developed [3]. When compared with HEVC, SHVC has even higher coding complexity, which adversely impacts its wide adoption. Therefore, it is very important to improve the coding speed of SHVC, especially for wireless and real-time video applications.

In this paper, we propose a fast depth and mode decision algorithm to accelerate the coding speed of intra prediction for quality SHVC (QS). The novelties and contributions of the proposed algorithm are summarized hereafter:

(1) Spatial and temporal correlations, as well as their corre-

lation degrees, are combined in a Naive Bayes classifier in prediction.

(2) Inter-Layer Reference (ILR) mode probability is combined with Partial Zero Blocks (PZBs) based on the Sum of Squared Differences (SSD) to determine the necessity of intra prediction checking.

(3) Initial Intra Modes (IMs) are obtained through Sobel operator, and residual coefficients are used to develop the relationship between two neighboring IMs in Hadamard Cost (HC) values for early termination in the Rate-Distortion Optimization (RDO) process.

(4) Depth probabilities are combined with the distribution of residual coefficients at the current depth to determine depth early termination.

Experimental results demonstrate that the proposed algorithm can significantly improve coding speed with negligible loss in coding efficiency.

The rest of this paper is organized as follows. Section 2 describes related work. Section 3 introduces the proposed framework. The four fast decision methods are then presented in details. Section 4 discusses the experimental results to demonstrate the performance of the proposed algorithm. Finally, Section 5 concludes this research and discusses future work.

II. RELATED WORK

Since there is a significant similarity among H.264/SVC, HEVC and SHVC, we first discuss fast algorithms for H.264/SVC, then review fast algorithms for HEVC, and finally we analyze fast algorithms for SHVC.

To improve the coding speed for H.264/SVC, several algorithms have been proposed. The authors of [4-6] use the coding mode of the co-located MB in BL to predict the current MB's candidate modes and skip unlikely modes in EL. Lee et al. [7] use motion vectors and All-Zero Blocks (AZB) of the co-located MB in BL to predict the current MB's likelihood modes in EL. Kim et al. [8] combine the co-located MB's mode in BL with RD cost of BLskip and Inter16×16 to predict the likelihood modes of the current MB in EL. Shen et al. [9] identify and utilize motion and mode characteristics of the current MBs, based on inter-layer and spatial correlations, to predict motion estimation step, mode decision, search-range selection, and direction selection. Shen et al. [10] adopt the SKIP modes of the co-located MB in BL and neighboring MBs in EL to predict the SKIP mode of the current MB in EL. The research in [11] statistically derives the Rate-Distortion (RD) cost expectation for each mode to encode. According to this expectation from small to large, it early terminates the encoding procedure when the RD cost is smaller than a pre-set threshold. Yeh et al. [12] use Bayesian theorem and Markov process to predict the best mode to reduce coding complexity. Jung et al. [13] use the co-located MB in BL and neighboring MBs to predict the probable AZBs, and then further examine and terminate these MBs by the AZB detection algorithm. Zhao et al. [14] use inter-layer and spatial correlations to sort the candidate mode list by likelihood in a descending order, and then early terminate the coding process based on a constrained model with optimal stopping. Lu et al. [15] predict candidate modes with early termination for the encoding process based on inter-layer and spatial correlations, textural features and motion activities. The work in [16] uses inter-layer and spatial correlations to predict candidate modes and stop the coding process early based on RD cost and residual coefficients. Generally speaking, these algorithms use correlations to separately predict candidate modes or early terminate the coding process, or exploit mode prediction and early termination together, to improve the coding speed.

Several fast algorithms have been proposed for HEVC in order to improve the coding speed. Zhao et al. [17] calculate gradient directions and obtain a gradient-mode histogram for each Coding Unit (CU). Based on the distribution of the histogram, only a small subset of candidate modes is selected in the RMD and RDO processes to improve the coding speed. Yan et al. [18] merge adjacent modes into groups based on the size of CU and HC values. Then, early termination and pixel-based edge detection methods are employed to further reduce the number of candidates for the RDO process. Zhang et al. [19] reduce IMs through the HC-based progressive rough mode search, and early terminate CU split according to the RD cost. Min et al. [20] calculate both global and local edge complexities of horizontal, vertical, 45 diagonal, and 135 diagonal directions and use them to decide whether to further divide a CU. In the same way, its four subCUs are also processed to decide whether to further divide them. To improve coding speed, Cho et al. [21] propose early CU split decision and early CU pruning decision at each CU depth level according to a Bayes decision rule method, based on low-complexity RD costs and full RD costs. Shen et al. [22] select likelihood depth levels based on spatial correlations, and then select likelihood prediction modes based on RD cost and prediction mode correlations among different depth levels or spatial correlations. Jamali et al. [23] predict candidate modes and skip some modes based on improved edge detection, neighboring blocks and classification of the Sum of Absolute Hadamard Transformed Difference (SATD) costs to improve the coding speed. In [24-26], Bayesian methods are used in intra prediction. Random forests via offline training [27] and Decision Tree [28] are used to predict the CU size of 64×64 or 32×32. In [29], linear SVM classifiers use depth difference and Hadamard/RD costs among spatial neighboring CUs to determine whether to early terminate the CU decision. Meanwhile, intra modes are predicted based on the gradients of horizontal and vertical directions. Liu et al. [30] use Convolutional Neural Network (CNN) to reduce the intra CU/PU modes.

Although both SVC and HEVC have significant similarities with SHVC, they also have some notable differences. Compared to SVC, SHVC uses some advanced coding tools, such as more CU sizes, more PU sizes and more intra modes. In addition, although both SHVC and HEVC use the same advanced coding tools, SHVC has a different architecture with a BL and inter-layer prediction. Consequently, fast algorithms developed for SVC and HEVC are often not applicable to SHVC. Therefore, it is crucial to improve the coding speed of SHVC based on its own features. In this paper, we mainly focus on coding speed improvement for QS.

Some techniques have been developed to improve the coding speed of SHVC. Tohidypour et al. [31-32] use the RD costs of the co-located block of the CU in BL and its four neighboring CUs in BL, and use the neighboring CUs' RD cost of the current CU in EL, to predict the current CU's RD cost in EL for early termination. In [33], the neighboring CUs' modes of the current CU in EL and the co-located CU's mode of the current CU in BL are used to predict the current CU's candidate modes in EL for quality SHVC [33]. Tohidypour et al. [34] use the depth and scalable layer of the current CU to predict candidate modes and exclude unlikely modes of the current CU in EL, and then exploit inter-layer and spatial correlations to further eliminate candidate modes to improve the coding speed. Based on spatial and inter-layer correlations, a probabilistic approach is proposed to predict the quad-tree structure of coding tree units (CTUs) for SHVC [35]. Tohidypour et al. [36] build a probabilistic model to predict the candidate modes of the current CU in EL to improve the coding speed based on relative CUs' coding information. Wang et al. [37] combine inter-layer correlation and the relationship of sizes between a CU's variance and the average variance of its four sub-CUs to determine whether to encode or skip some depths. Then, they select directional modes of relative CUs, and the directional mode with the smallest HC cost in mode 2, 10, 18, 26, and 34, as initial modes, and finally combine with the relationship between the modes and HC to obtain the best directional mode. Wang et al. [38] first predict candidate depths, then predict the best mode between ILR mode and intra mode in each candidate depth, and finally test whether depth early termination is satisfied to skip other depths to improve the coding speed.

Although the above algorithms can improve the coding speed, they have not considered some aspects to further speed up the coding process, including: (1) combining spatial and inter-layer correlations usually uses Bayes classifier in prediction, however, their degree of correlation has not been fully considered, which might affect the accuracy of prediction; (2) AZBs are usually developed for early termination, but the number of AZBs is limited, thus only exploiting AZBs restricts the speed up gain. In addition, AZBs are developed based on the Sum of Absolute Difference (SAD). Employing PZBs based on SSD to estimate whether the ILR mode is the best mode can significantly improve the coding speed; (3) Sobel operator is often used to predict intra modes, and HC values are also used to predict IMs. However, both Sobel operator and HC values are not fully and simultaneously exploited, which might cause insignificant improvement in coding speed; (4) If the prediction is very accurate, residual coefficients should obey a Laplacian distribution. In addition, depth probabilities also influence depth selection. Combining depth probabilities and the distributions of residual coefficients can significantly improve the coding speed.

Based on the above observations, in this paper, we propose a new and fast intra prediction algorithm for QS. First, spatial and temporal correlation as well as their correlation degrees are jointly used to calculate depth probabilities and skip depths with low likelihoods. Then, the ILR mode probability is combined with SSD based PZB to determine the necessity of intra prediction checking, where RMD and RDO procedures are included. In the RMD procedure, Sobel operator is combined with the relationship between IMs and their corresponding HC values to predict candidate IMs. The residuals of two neighboring IMs are used to derive the early termination condition for RDO procedure. Finally, the residual coefficients of the current depth and its probability are tested to determine early termination.

III. PROPOSED FAST INTRA PREDICTION PROCESS

In order to improve intra coding speed and maintain coding efficiency for QS, four strategies are proposed: Correlation-Based Depth Prediction (CBDP); Mode Probability and Partial Zero Block-Based ILR Mode Decision (MPPZBB-IMD); Gradient and Hadamard-Cost Based intra Mode Prediction (GHCB-IMP); and Depth Probability and Residual Distribution Based Depth Early Termination (DPRDB-DET). The right side of Fig.1 illustrates the flowchart of the proposed algorithm, while the left side show the four strategies. First, CBDP is used to obtain depth probabilities and skip depths with low likelihood. For the selected current depth candidates, MPPZBB-IMD determines whether the best mode is the ILR mode. In the affirmative, intra prediction can be directly skipped without the need of checking; otherwise the IM candidates are predicted through GHCB-IMP. After the depth has been checked, the best depth is predicted by DPRDB-DET for early termination.

In order to develop the four strategies, the relative methods and relationships in intra coding are investigated based on extensive experiments. To meet the diverse requirements for video resolutions, we use test sequences in class B, C, D and E in our experiments, and select two test sequences in each class: Sunflower and Tractor in class B; Flowervase and PartyScene in class C; BlowingBubbles and RaceHorses in class D; and Park and Town in class E. Since we study intra only algorithm for QS, each test sequence is encoded using all I-frame structure. According to common SHM test conditions (CSTC) [39], the QPs used for the BL are set as (26, 30, 34, 38), and the corresponding QPs used for the EL are set as (22, 26, 30, 34) and (20, 24, 28, 32), respectively. Since the performances of the two QP settings are similar, only the experimental results for QPs with (22, 26, 30, 34) are provided for brevity. Based on these experiments, we propose our efficient fast intra prediction methods as discussed below.

A. Correlation-Based Depth Prediction (CBDP)

Using a quadtree structure in SHVC, a CU is divided from depth 0 to 3, with a corresponding CU size from 64×64 to 8×8. For every depth, an encoder has to perform the whole intra prediction and ILR prediction process. Therefore, SHVC needs to encode four depths, which include four rounds of the entire intra and ILR prediction processes. It is straightforward that skipping depths with low likelihood can significantly improve the coding speed. Since neighboring CUs are usually very similar, their spatial correlation is very high. The current and

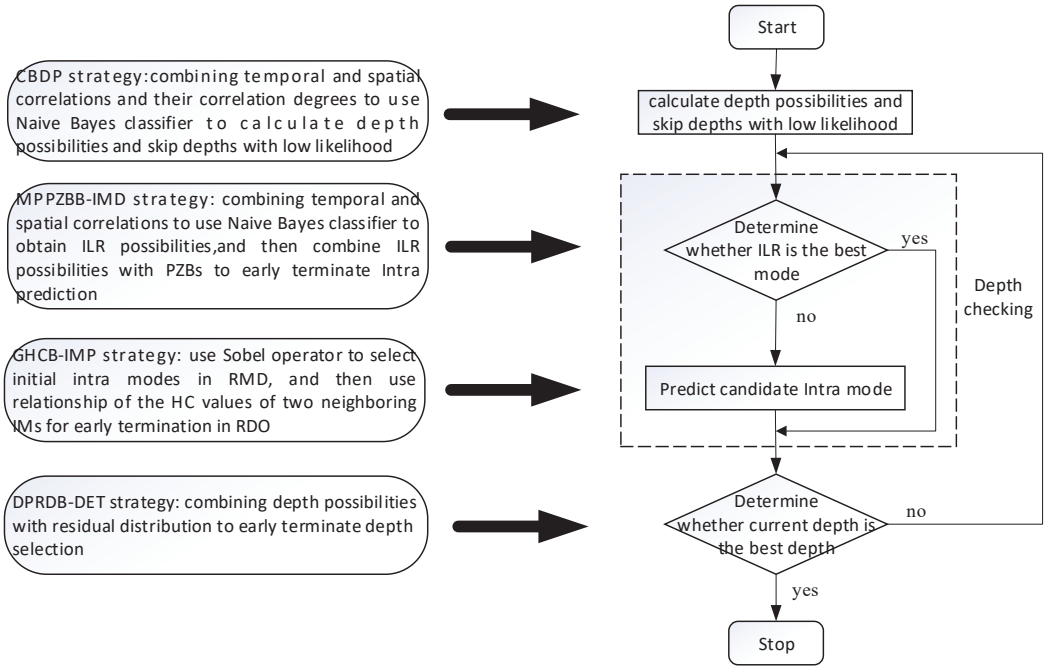

Fig. 1: Flowchart of the overall proposed algorithm

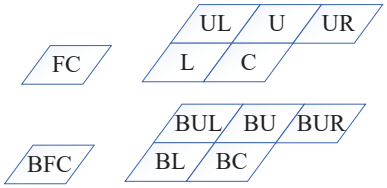

Fig. 2: Neighboring CUs of current CU in depth prediction.

previous frames have also high temporal correlation. Therefore, we can use spatial and temporal correlations to predict probable depths and skip unlikely ones, hence improving the coding speed [40-41]. The spatial and temporal correlations are exploited through the neighboring CUs of the current CU, as shown in Fig.2. The current CU in the EL, its co-located CU in the previous frame, its left CU, upper CU, upper-left CU and upper-right CU are denoted as C, FC, L, U, UL and UR, respectively. Accordingly, the collocated CUs of C, FC, L, U, UL and UR in the BL are denoted as BC, BFC, BL, BU, BUL and BUR, respectively.

As we know, since spatial and temporal correlations are present, spatial and temporal neighboring CUs' depths are usually used to predict the current CU's coding depth candidates. As different neighboring CUs may have different degrees of correlation with the current CU, using neighboring CUs in prediction without considering this characteristic may not obtain the best performance. Therefore, we should combine neighboring CUs and their corresponding degrees of correlation to calculate the depth probability of current depth through the following Bayes rule. Assume cd is one of the depth levels of the current CU in EL, **nd** and **nr** are the depth level vector and correlation degree vector of neighboring CUs of the current CU in EL. In this condition, the corresponding probability of the current CU using depth level cd, $f_d(cd)$, is derived as

$$f_d(cd) = p\left(cd \middle| (nd, nr)\right) = \frac{p\left((nd, nr), cd\right)}{p\left(nd, nr\right)} = \frac{p\left((nd, nr) \middle| cd\right) p\left(cd\right)}{p\left(nd, nr\right)}, \tag{1}$$

where $p(cd)$ is the probability of the current CU using depth level cd, $p(nd, nr)$ is the probability of neighboring CUs with a vector **(nd, nr)**, $p(nd, nr|cd)$ is the conditional probability of neighboring CUs with a vector **(nd, nr)** given the current CU using depth level cd. Since both CUs in the BL and the co-located CUs in the EL are the same except for QPs, the degrees of correlation of CUs in the EL can be set as equal to the ones of the co-located CUs in the BL. Obviously, when the absolute difference of neighboring CUs in depth is smaller, their spatial correlation is stronger; and vice versa. In other words, the relationship between the degrees of correlation and the absolute differences of neighboring CUs in depth in the BL is inversely proportional. Suppose **nbd** is the depth level vector of neighboring CUs of the CU BC in BL. Since the maximum absolute difference of neighboring CUs in depth is 3, the i-th ($0 \le i \le 4$) component of the correlation degree vector, nr_i, can be written as

$$nr_i = 3 - \left|nd_i - nbd_i\right|, \tag{2}$$

where nd_i and nbd_i refer to the i-th ($0 \le i \le 4$) component of the depth level vector **nd** and **nbd**, respectively.

Since the current CU has five neighboring CUs, each vector has five components, and each component takes four values, namely 0, 1, 2 and 3. If we directly use Eq. (1) in calculation, the process is very complex. In order to overcome this issue, we can use the Naive Bayes classifier which can make a conditional independence assumption. In other words, we

TABLE I: Depth probability distribution $p(nd_i, nr_i)$

nd_i \ nr_i	0	1	2	3
0	8.1%	6%	2.3%	3.6%
1	4.7%	2.7%	1.4%	0.7%
2	13.3%	1.7%	1.1%	0.6%
3	47.6%	3%	2.1%	1.1%

assume that the depth and degree of correlation of each CU is independent of each other. That is to say, different components of a vector are independent. According to this independence assumption, Eq.(1) can be computed as follows:

$$f_d(cd) = \frac{p((nd,nr)|cd)\,p(cd)}{p(nd,nr)} = \frac{p(cd)\prod_{i=0}^{4} p((nd_i,nr_i)|cd)}{\prod_{i=0}^{4} p(nd_i,nr_i)}. \tag{3}$$

Since these depth probability distributions in different positions have small differences, for the convenience of calculation, we use their average values instead. For example, $p(0,0)$ in CU C, FC, L, U, UL and UR are set as their average values. Therefore, these depth probability distributions have no relationship with their positions. In other words, different components should have the same depth probability distribution. Therefore, we only need to list the probability distribution of one component. According to the experimental conditions mentioned above, we can obtain the depth probability distributions of the i-th ($0 \le i \le 4$) component in the vector **(nd, nr)** and **(nd, nr|cd)**, which are respectively denoted as $p(nd_i, nr_i)$ and $p((nd_i, nr_i)|cd)$. They are listed in Table I and Table II, respectively.

From the experiments, $p(cd)$ in depth level 0, 1, 2 and 3 are 16.6%, 9.2%, 15.4% and 58.8%, respectively. The conditional probability of the current CU using depth level cd, $f_d(cd)$, can be obtained according to Eq. (3). Since the computation may involve some rounding errors, the sum of probabilities of four depth levels may not always be equal to 1, we can amend it as follows

$$f_d(cd) = \frac{f_d(cd)}{f_d(0) + f_d(1) + f_d(2) + f_d(3)}. \tag{4}$$

In this way, we can obtain the probabilities of all four depth levels. The probability threshold can be set to 0.95 according to [35]. If the probability is larger than or equal to 0.95, the current CU will almost certainly select the corresponding depths, and the other depths are very unlikely. Based on the above analysis, we propose our depth selection method below:

(1) If the probability of depth level 0 is smaller than 0.05, the depth level is very unlikely to be selected, thus depth level 0 can be directly skipped.
(2) If the sum of probabilities of depth level 0 and 1 are smaller than 0.05, these depth levels are very unlikely to be selected and can be directly skipped.
(3) If the sum of depth probabilities from depth level 0 to the current depth is larger than or equal to 0.95, the current CU almost certainly use one of these depth levels. Thus, we can skip subsequent depth levels.
(4) If the depth 3 needs to be checked, extensive experiments show that 4×4 CUs are seldom selected. Thus 4×4 CUs are directly skipped.

In order to demonstrate the effectiveness of the proposed depth selection method, the corresponding performance is listed in Table III, in which BDBR [42] measures the bitrate difference compared with the SHVC reference software (SHM 11.0) at an equal PSNR in the EL. A positive BDBR indicates a decrease in coding efficiency, while a negative value indicates an improvement. TS denotes the percentage of encoding run-time savings in the EL only.

In Table III, we observe that the average BDBR is 0.44% and the average TS is 52.16%. Therefore, the proposed depth selection method can significantly improve the coding speed with negligible losses in coding efficiency.

B. Mode probability and Partial Zero Block-Based ILR Mode Decision (MPPZBB-IMD)

In QS, since frame resolutions between the BL and EL are the same, the inter-layer correlation is very high. Therefore, a CU in the EL searches for the best matching CU in the reconstructed pixels in BL, by using the ILR mode. It is usually an accurate prediction and many CUs may select it as the best mode. Nevertheless, there is a low, but non-negligible, probability that intra mode is optimal [38]. If we always check intra mode, unnecessary time will be spent. However, if intra mode is completely skipped, coding efficiency will be obviously degraded. In this condition, we can check the ILR mode, and then determine whether it is the best mode. In the affirmative, intra mode can be directly skipped. Otherwise, intra mode needs to be further tested. The key problem is how to determine if the ILR mode is the best one. In general, if a mode is a good predictor, the residue is usually very small, with many quantized residual coefficients being zero. Therefore, AZBs are usually used to predict early termination, as in [13]. However, since the number of AZBs is limited, when only using AZBs in prediction, the coding speed improvement is limited. In order to further improve coding speed, we do not need to use the above strict condition. In practice, the percentage of zero quantized coefficients in a CU could also be an indicator of the best prediction mode. Therefore, we can use PZBs to early terminate intra prediction.

PZBs can be calculated based on the distribution of residual coefficients. Residual coefficients are typically modeled using a Laplacian distribution [43]. In SHVC, there are four different CUs with 64×64, 32×32, 16×16 and 8×8 sizes. We can calculate the PZB of 8×8 CU, and then calculate the other three CUs' PZBs. Since different CUs may have different probabilities to use ILR mode, using a uniform condition for all CUs cannot achieve the best performance. Therefore, we should combine probabilities with PZBs to set early termination condition of ILR.

In addition, AZBs (PZBs) are generally predicted based on SAD of residual coefficients. Obviously, SSD of residual coefficients is more accurate than SAD to predict PZBs. Therefore, we first obtain the probability of using ILR of a CU, then we derive PZBs based on square roots of SSD of

TABLE II: Conditional depth probability distribution $p((nd_i, nr_i)|cd)$

nr_i \ cd / nd_i	0				1				2				3			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
0	20.6%	11.9%	3.3%	1.1%	10.8%	8.5%	2.5%	1.5%	5.3%	3.9%	2.7%	3.5%	3%	2%	2%	6%
1	5.3%	3.4%	1.5%	1.2%	12.5%	5.5%	1.9%	0.7%	6%	2.8%	1.8%	0.9%	2.8%	1.4%	1.2%	0.7%
2	11.4%	2.5%	1.4%	1%	15.5%	2.5%	1.5%	1%	27.4%	3%	1.7%	0.8%	10.8%	1.2%	0.8%	0.6%
3	27%	4.2%	2.4%	1.8%	28.9%	3.2%	2.1%	1.4%	34.7%	3%	1.7%	0.8%	61%	3.2%	2.2%	1.1%

TABLE III: The performances for the CBDP

Sequence	BDBR	TS
Sunflower	0.1%	55.05%
Tractor	0.1%	49.8%
Flowervase	0.5%	54.65%
PartyScene	0.3%	53.35%
BlowingBubble	0.5%	52.08%
RaceHorses	0.8%	53.3%
Parkrunner	0.2%	45.03%
Town	1.0%	54.03%
Average	0.44%	52.16%

residual coefficients, and finally we combine probabilities with PZBs to develop early termination condition of ILR.

1) Correlation-Based Mode Prediction (CBMP): In QS, the BL only encodes in intra mode, whereas the EL uses both intra and ILR modes. Since intra mode in BL cannot provide any information to predict ILR mode in EL, we therefore only use relative CUs in EL to predict ILR mode in EL. The neighboring CUs in EL are the same as in Fig.2. As mentioned above, exploiting both spatial and temporal correlations, we can use neighboring CUs in EL to predict the probability of the current CU using mode cm, $f_m(cm)$, by the following Bayes rule

$$f_m(cm) = p(cm|nm) = \frac{p(nm, cm)}{p(nm)} = \frac{p(nm|cm)\,p(cm)}{p(nm)}, \quad (5)$$

where **nm** is the mode vector of neighboring CUs of the current CU in EL. Similar to depth prediction, we can use the Naive Bayes classifier. We can assume that the mode of each CU is independent of each other. In other words, the different components of **nm** are independent. According to this independence assumption, Eq. (5) can be rewritten as follows

$$f_m(cm) = \frac{p(nm|cm)\,p(cm)}{p(nm)} = \frac{\prod_{i=0}^{4} p(nm_i|cm)\,p(cm)}{\prod_{i=0}^{4} p(nm_i)}. \quad (6)$$

Similar to depth prediction, since mode probability distribution in different positions have small differences, we also use their average value. Therefore, different components should have the same mode probability distribution. We only need to list the probability distribution of one component. According to the experimental conditions mentioned above, we can obtain the conditional mode probability distribution of the i-th ($0 \le i \le 4$) component in the vector **nm|cm**, $p(nm_i|cm)$, which is listed in Table IV.

TABLE IV: Conditional mode probability distribution $p(nm_i|cm)$

nm_i / cm	ILR	intra
ILR	83.6%	16.4%
intra	73.2%	26.8%

Since the computation may involve some rounding errors, the sum of probabilities of two modes may not always be equal to 1, we can amend it as follows

$$f_m(cm) = \frac{f_m(cm)}{f_m(m_{ILR}) + f_m(m_{\mathrm{intra}})}, \quad (7)$$

where m_{ILR} and m_{intra} refer to ILR and intra modes, respectively. In this way, we can obtain the probabilities of both modes. We also use 0.95 as the mode probability threshold [35]. If the probability of ILR is larger than or equal to 0.95, the current CU only uses ILR mode to encode. We evenly divide mode probabilities into three classes: (1) high mode probability (probability is larger than or equal to 0.64); (2) medium mode probability (probability is larger than or equal to 0.32); (3) low mode probability (probability is smaller than 0.32).

2) Partial Zero Block-Based SSD (PZBB-SSD): Supposing an 8×8 residual block $x(i, j)$, ($0 \le i, j \le 7$), are approximated by a Laplacian distribution with zero mean and variances σ^2 [43], the PZBs based on SSD are then calculated as described below.

The expected value of $x(i, j)^2$ is given by

$$E\left(x(i,j)^2\right) = D(x(i,j)) - (E(x(i,j)))^2, \quad (8)$$

where $E(x(i, j))$ is the expected value of $x(i, j)$, $D(x(i, j))$ is the variance of $x(i, j)$. Since $E(x(i, j))$ =0 [43], we can derive

$$E\left(x(i,j)^2\right) = D(x(i,j)) = \sigma^2. \quad (9)$$

From probability theory, the average approximates the expected value, and then

$$\sigma^2 = \frac{SSD}{64}, \quad (10)$$

where SSD is the Sum of Squared Differences of the residual coefficients for an 8×8 block. The variance of the $(i, j)^{th}$ DCT coefficient, $\sigma_F^2(i, j)$, can be written as [43]

$$\sigma_F^2(i,j) = \sigma^2 \left[ARA^T\right]_{i,i} \left[ARA^T\right]_{j,j}, \quad (11)$$

$$\begin{bmatrix} 0.7219 & 0.9439 & 1.2095 & 1.5480 & 1.8720 & 2.1496 & 2.3606 & 2.4902 \\ 0.9439 & 1.2342 & 1.5814 & 2.0240 & 2.4476 & 2.8105 & 3.0864 & 3.2560 \\ 1.2095 & 1.5814 & 2.0265 & 2.5935 & 3.1364 & 3.6014 & 3.9550 & 4.1722 \\ 1.5480 & 2.0240 & 2.5935 & 3.3192 & 4.0140 & 4.6092 & 5.0616 & 5.3397 \\ 1.8720 & 2.4476 & 3.1364 & 4.0140 & 4.8541 & 5.5739 & 6.1211 & 6.4573 \\ 2.1496 & 2.8105 & 3.6014 & 4.6092 & 5.5739 & 6.4004 & 7.0287 & 7.4148 \\ 2.3606 & 3.0864 & 3.9550 & 5.0616 & 6.1211 & 7.0287 & 7.7187 & 8.1427 \\ 2.4902 & 3.2560 & 4.1722 & 5.3397 & 6.4573 & 7.4148 & 8.1427 & 8.5900 \end{bmatrix}. \tag{15}$$

where $[.]_{i,i}$ is the $(i,i)^{th}$ component of a matrix, A is an 8×8 integer DCT transform divided by $128\sqrt{2}$ in SHVC, R is

$$R = \begin{bmatrix} 1 & \rho & \cdots & \rho^7 \\ \rho & 1 & \cdots & \rho^6 \\ \vdots & \vdots & \vdots & \vdots \\ \rho^7 & \rho^6 & \cdots & 1 \end{bmatrix}, \tag{12}$$

where ρ is the correlation coefficient, which is set to 0.6 [43]. For a zero-mean Laplacian distribution, the probability of a value falling within $(-3\sigma, 3\sigma)$ is about 99.73%. If the value is quantized to zero, it must satisfy the following

$$3\sigma_F(i,j) < 5Q_{step}/6. \tag{13}$$

Combining Eqs. (10), (11) and (13), the following expression is obtained

$$\sqrt{SSD} \le 20Q_{step} \Big/ \left(9\sqrt{\left[ARA^T\right]_{i,i}\left[ARA^T\right]_{j,j}}\right). \tag{14}$$

The right-hand coefficient value of Eq. (14) corresponding to the threshold matrix is given in Eq. (15).

After the threshold matrix is obtained, we should select a threshold value in the matrix based on mode probability.

3) ILR mode early termination based on mode probabilities and PZBs (IMETB-MPP): We have divided mode probabilities into high, medium and low mode probabilities. We should select the best threshold value in the matrix based on mode probabilities, and then the corresponding PZB can be obtained. For example, if $\sqrt{SSD} < 0.7219Q_{step}$, the whole quantized residual coefficients are zero; if $\sqrt{SSD} < 0.9439Q_{step}$, only the quantized residual coefficient in (0, 0) is not zero; if $\sqrt{SSD} < 1.2095Q_{step}$, only quantized residual coefficients in (0, 0), (0, 1) and (1, 0) are not zero; and so on. In the same way, we can obtain other PZBs. If we select too small values to test, the coding speed improvement will remain limited. On the other hand, if we select too large values to test, the coding efficiency will be degraded significantly.

For ILR with high probabilities, we can select some larger values in the matrix to test. In order to select the best threshold value, we have therefore used 2.3606, 2.4902, 3.256, 4.1722, 5.3397 and 6.4573 in our experiments.

From Table V, we can observe that when the threshold value is larger than 4.1722, sequence "flowervase" has obviously changed in BDBR. Therefore, we select 4.1722 as the best threshold value. In the same way, for ILR with medium probabilities, the best threshold value is set to 2.4902. For depths with low probabilities, the best threshold value is set to 2.1496.

Through the above process, the best threshold values for ILR with high, medium and low probabilities are 4.1722, 2.4902 and 2.1496, respectively. The threshold values are only developed for 8×8 CUs, we can obtain threshold values for other size CUs based on 8×8 CUs. Since each 64×64, 32×32, 16×16 CU includes respectively 64, 16 and 4 8×8 CUs, the corresponding thresholds for 64×64, 32×32, 16×16 and 8×8 CU can be written as

$$\sqrt{SSD} \le \begin{cases} 4.1722mQ_{step} & \text{high} & \text{probability} \\ 2.4902mQ_{step} & \text{medium} & \text{probability,} \\ 2.1496mQ_{step} & \text{low} & \text{probability} \end{cases} \tag{16}$$

where m is the number of 8×8 CUs included and is 64, 16, 4 and 1 for 64×64, 32×32, 16×16 and 8×8 CUs, respectively. It is worth noting that there are some sequences with negative BDBR values, which suggests that the proposed scheme obtains small BD rate savings for some sequences. Possible reasons will be discussed at the end of Section 4.

C. Gradient and Hadamard-Cost Based Intra Mode Prediction (GHCB-IMP)

Similar to HEVC, there are 35 IMs including DC and Planar modes in SHVC. During the RMD process, all the 35 IMs are checked to select the first N IMs with the smallest SATD costs. N takes different values according to the corresponding CU size. If the size of a CU is equal or greater than 16×16, N is 3; otherwise N is 8. Afterwards, the IMs of the left and upper partition units (PUs) in EL and the co-located PUs in BL are further selected as candidate modes. Then, the best IM is obtained among those candidates in the RDO process.

From the above description, we can see that the intra prediction for SHVC employs both RMD and RDO processes. In order to improve the coding speed, we propose the below new methods to speed up these two processes respectively.

1) Improved Gradient for Rough Mode Prediction: Gradient has been widely used in prediction for intra modes [17] [23]. Due to the diverse video sequences, it is very difficult to obtain the optimal number of gradients. If more gradients are selected, coding speed will not be significantly improved. On the contrary, if few gradients are selected, coding efficiency will be obviously degraded. To overcome the problem, we propose to first obtain the most probable gradients, and then integrate the relationship between IMs and their corresponding HC values to predict the first N IMs. The details are discussed below.

By using a classical Sobel operator with two 3×3 convolution masks, we can compute both components of a gradient for a pixel

TABLE V: Threshold values and the corresponding BDBR

Sequence \ threshold values	2.3606	2.4902	3.256	4.1722	5.3397	6.4573
Sunflower	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Tractor	-0.10%	-0.10%	-0.10%	-0.10%	-0.10%	-0.10%
Flowervase	-0.20%	-0.30%	-0.30%	-0.30%	-0.50%	-0.50%
PartyScene	0.00%	0.00%	-0.10%	-0.10%	-0.10%	-0.10%
BlowingBubble	0.00%	-0.10%	0.00%	-0.10%	-0.10%	-0.10%
RaceHorses	-0.10%	-0.10%	-0.10%	-0.10%	-0.10%	-0.10%
Parkrunner	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Town	-0.10%	-0.10%	-0.20%	-0.20%	-0.20%	-0.20%

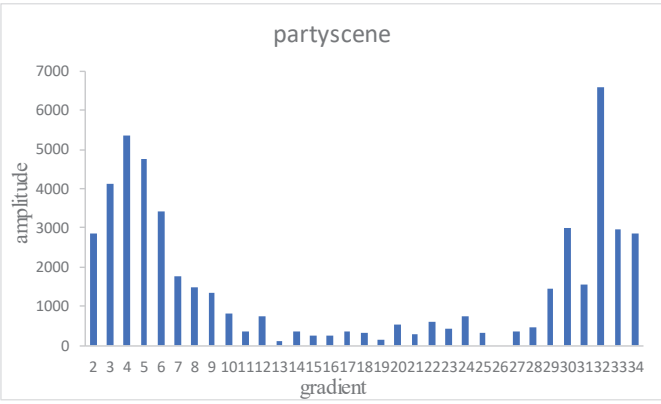

Fig. 3: Gradient Histogram for PU size 32×32

$$\begin{aligned} G^x_{i,j} &= p_{i-1,j+1} + 2 \times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - \\ &\qquad 2 \times p_{i,j-1} - p_{i+1,j-1} \\ G^y_{i,j} &= p_{i+1,j-1} + 2 \times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - \\ &\qquad 2 \times p_{i-1,j} - p_{i-1,j+1}, \end{aligned} \tag{17}$$

where $p_{i,j}$ is the pixel value at (i,j), $G^x_{i,j}$ and $G^y_{i,j}$ represent the vertical and horizontal components of a gradient, respectively. We can calculate the amplitude of the gradient by

$$Amp\left(G_{i,j}\right) = \left|G^x_{i,j}\right| + \left|G^y_{i,j}\right|. \tag{18}$$

The direction of the gradient is obtained according to

$$Ang\left(G_{i,j}\right) = \arctan\left(\frac{G^y_{i,j}}{G^x_{i,j}}\right). \tag{19}$$

In order to simplify calculation, $G^y_{i,j} \Big/ G^x_{i,j}$ is actually used instead of arctan for optimization.

From the above process, the gradient of each pixel can be obtained, and then the amplitudes of similar gradients are summed up. According to the IMs in SHVC, we can calculate the limit of any IM in Table VI. If the directions of gradients are within the same scope, these gradients are similar. Since modes 0 and 1 are non-directional, only 33 directional modes are listed in Table VI.

We can obtain a gradient histogram by using the above method. Fig.3 illustrates a gradient histogram for PU size 32×32 for the "PartyScene" sequence. Obviously, gradients with greater larger sums have larger probabilities of being the best intra mode. However, gradients obtained by Sobel operator are different from IMs obtained by Hadamard transform. Therefore, it is very difficult to accurately predict the best IM only through gradients. We propose to integrate gradients with the relationships between IMs and their corresponding HC values to obtain the first N IMs.

TABLE VI: Upper & lower limits of $G^y_{i,j}/G^x_{i,j}$ for intra modes

mode	lower	upper	mode	lower	upper
2	-1.10879	-0.90189	19	1.108785	1.367109
3	-0.90189	-0.73147	20	1.367109	1.689078
4	-0.73147	-0.59204	21	1.689078	2.140169
5	-0.59204	-0.46725	22	2.140169	2.919258
6	-0.46725	-0.34255	23	2.919258	4.588474
7	-0.34255	-0.21794	24	4.588474	9.162712
8	-0.21794	-0.10914	25	9.162712	32.03108
9	-0.10914	-0.03122	26	32.03108	$+\infty$
10	-0.03122	0.03122	26	$-\infty$	-32.0311
11	0.03122	0.109138	27	-32.0311	-9.16271
12	0.109138	0.217937	28	-9.16271	-4.58847
13	0.217937	0.342553	29	-4.58847	-2.91926
14	0.342553	0.467253	30	-2.91926	-2.14017
15	0.467253	0.592039	31	-2.14017	-1.68908
16	0.592039	0.731471	32	-1.68908	-1.36711
17	0.731471	0.901888	33	-1.36711	-1.10879
18	0.901888	1.108785	34	-1.10879	-0.90189

Fig. 4 shows the relationship between IMs and their corresponding HC values for "PartyScene" sequence. For different sequences, we can divide 35 IMs into different numbers of intervals. In Fig. 4, the horizontal axis represents IMs and the vertical axis corresponds to HC values. Within each interval, both the left and right neighboring IMs monotonically approach to their local minimum points (LMPs) along with the decreasing of HC values [37-38]. There are five zones A, B, C, D and E, and each zone contains a LMP. Since RD costs have very strong correlations with HC values, it is highly possible that the best IM is included in the LMPs and their neighbors. Therefore, we should definitely search the LMPs. According to the above observation, by following the descent direction of the HC values, we can obtain the LMP within each zone. Here, how to select the best initial IMs is the key point.

Since gradients with larger sums have higher probabilities of being the best IM, we can choose some gradients with the highest sums as the initial IMs, and then follow the descent

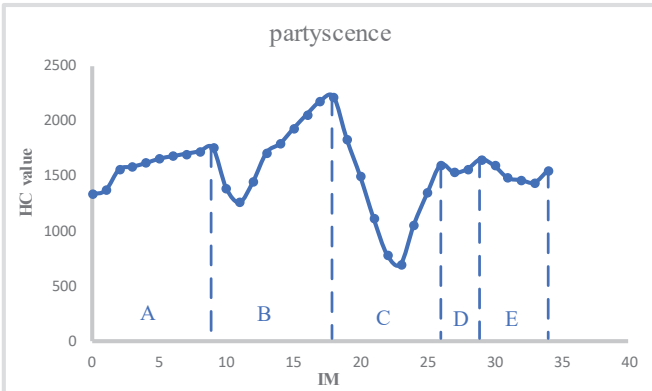

Fig. 4: The relationship between IMs and HC values

TABLE VII: BDBR corresponding to the number of gradients selected

Sequence	BDBR		
	One	Two	Three
Sunflower	0.1	0	0
Tractor	0	0	0
Flowervase	-0.1	-0.1	-0.1
PartyScene	0	0	0
BlowingBubbles	0	0	0
RaceHorses	0	0	0
Park	0	0	0
Town	0	0	0
Average	0	-0.013	-0.013

direction of the HC values to search the first N modes. The selection of the number of gradients with the highest sums is very crucial. We have conducted experiments to investigate the best number of gradients to be selected.

For large CUs, such as 64×64, 32×32, 16×16, since 3 candidate modes need to be checked in the RDO process, we have selected the first three gradients as initial IMs to search. According to our experimental results, the corresponding BDBRs are listed in Table VII, in which "One", "Two" and "Three" represent the number of gradients selected.

From Table VII, we can observe that selecting two gradients gives the same coding efficiency as three gradients. It is very straightforward that the fewer gradients selected, the more the coding speed is improved. As a result, we adopt the first two gradients in our experiments. Using the same approach, we have tested CUs with 8×8 and 4×4. Our experiments also demonstrate that selecting the first two gradients as initial IMs achieves the best performance.

For the usage of the HC, since we only select two initial IMs and then follow the descent direction of the HC values to search the LMP within each zone, many unnecessary IMs are skipped in prediction and much unnecessary coding time can be saved. From Table VII, we can observe that the number of gradients selected has a very marginal impact on performance. The reason is that intra mode represents only a small percentage, no more than 30%, as shown in Table IV.

2) Hadamard-Cost Based Early Termination for RDO: After the RMD process, large CUs (64×64, 32×32, 16×16) with 3 IMs and small CUs (8×8, 4×4) with 8 IMs will be further tested in the RDO process to select the best IM. In order to improve the coding speed, we investigate the Hadamard costs of neighboring intra modes and propose to exploit them to make the decision for early termination. Suppose R_1 and R_2 are the CU residuals of two neighboring IMs, their difference R is

$$R = R_1 - R_2. \tag{20}$$

Through Hadamard transform, we have

$$HRH = HR_1H - HR_2H, \tag{21}$$

where H is a Hadamard matrix. Based on Cauchy-inequality, we obtain

$$\begin{aligned} HRH &\leq \left|\sum_{i=0}^{m}\sum_{j=0}^{m}\left(HH^T\right)^2\right|^{\frac{1}{2}} \times \left|\sum_{i=0}^{m}\sum_{j=0}^{m} r^2(i,j)\right|^{\frac{1}{2}} \\ &\leq \sqrt{m}\left|\sum_{i=0}^{m}\sum_{j=0}^{m} r^2(i,j)\right|^{\frac{1}{2}}, \end{aligned} \tag{22}$$

where m is the corresponding CU size. Through inequality (22), we derive

$$\sqrt{m}\left|\sum_{i=0}^{m}\sum_{j=0}^{m} r^2(i,j)\right|^{\frac{1}{2}} \leq \sqrt{m}\sum_{i=0}^{m}\sum_{j=0}^{m}|r(i,j)|. \tag{23}$$

Any value of (i,j) in HRH is given by

$$x_{i,j} = \sum_{k=0}^{m}\sum_{p=0}^{m} h_{ik} r_{kp} h_{pj} \leq \sum_{k=0}^{m}\sum_{p=0}^{m}\left|h_{ik}h_{pj}\right|\left|r_{kp}\right| \leq \sum_{k=0}^{m}\sum_{p=0}^{m}\left|r_{kp}\right|. \tag{24}$$

If R_1 and R_2 have no significant differences, any quantized values of HRH should be equal to 0. Then we can derive

$$\sum_{k=0}^{m}\sum_{p=0}^{m}\left|r_{kp}\right| < Q_{step}. \tag{25}$$

Combining Eq. (21), (22), (23), (25), we finally obtain

$$|HR_1H - HR_2H| < \sqrt{m}Q_{step}. \tag{26}$$

Eq.(26) can be written as

$$|HC_1 - HC_2| < \sqrt{m}Q_{step}, \tag{27}$$

where HC_1 and HC_2 refers to Hadamard transform values of two neighboring IMs. In other words, if two neighboring IMs satisfy the condition in Eq. (27), we can assume that they have no significant difference. However, this is only the theoretical condition. We can further adjust experimentally the right side of Eq. (27) to further improve the coding speed. More precisely, we divide the right side of Eq. (27) by 1, 2, 4, 8, 16 and 32. Based on experimental results, we can adjust the condition as follows.

$$|HC_1 - HC_2| < \sqrt{m}Q_{step}\Big/k. \tag{28}$$

When two neighboring IMs are the first two intra modes, $k = 4$; when the neighboring modes are the second mode and the third mode, $k = 8$; when the neighboring modes are the third mode and the fourth mode (only for CU with 8×8 and 4×4), $k = 32$. Since the other modes have low probabilities to be selected as the best mode, we have not further tested them in our experiments.

D. Depth Probability and Residual Distribution Based Depth Early Termination (DPRDB-DET)

In the above-described process, both spatial and temporal correlations are used to predict candidate depths and skip depths with low probabilities, in order to improve the coding speed and maintain the coding efficiency. As mentioned above, if a depth is predicted accurately, residual coefficients will obey a Laplacian distribution [43]. Therefore, the distribution of residual coefficients can be checked for early termination. For this purpose, we propose an experiment based on χ^2 fitting test. Since depth early termination is only used for depth 0, 1 and 2, the corresponding CU sizes are 64×64, 32×32, and 16×16, respectively. In order to test whether the residual coefficients obey a Laplacian distribution, we evenly divide residual coefficients into 16 zones, from the minimum value -255 to the maximum one 255, in increasing order. Suppose $x_1, x_2, \cdots, x_n$ are residual coefficients, with n the number of residual coefficients, the probability density function of the Laplacian distribution is

$$f(x) = \frac{1}{2\lambda}e^{-\frac{|x-\mu|}{\lambda}}. \tag{29}$$

The corresponding expected value $E(x)$ and variance $D(x)$ are

$$E(x) = \frac{1}{2\lambda}\int_{-\infty}^{\infty} xe^{-\frac{|x-\mu|}{\lambda}}dx = \mu, \tag{30}$$

$$D(x) = E(x - E(x))^2 = \frac{1}{2\lambda}\int_{-\infty}^{+\infty}(x-\mu)^2 e^{-\frac{|x-\mu|}{\lambda}}dx = 2\lambda^2. \tag{31}$$

The corresponding expected value $E(x)$ and variance $D(x)$ can also be calculated by

$$E(x) = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{32}$$

$$D(x) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}. \tag{33}$$

Combining Eqs. (30), (31), (32) and (33), we derive

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}, \lambda = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{2(n-1)}}. \tag{34}$$

Let us define the maximum value, minimum value and the number of residual coefficients in the i-th zone as a_i, b_i and m_i respectively. After μ and λ have been obtained in Eq. (34), the probability p_i of the i-th zone is given by

$$p_i = \begin{cases} \frac{e^{\frac{\mu-a_i}{\lambda}} - e^{\frac{\mu-b_i}{\lambda}}}{2} & a_i > \mu \\ \frac{e^{\frac{b_i-\mu}{\lambda}} - e^{\frac{a_i-\mu}{\lambda}}}{2} & b_i < \mu \\ \frac{2 - e^{\frac{a_i-\mu}{\lambda}} - e^{\frac{\mu-b_i}{\lambda}}}{2} & a_i \le \mu \le b_i \end{cases} \tag{35}$$

According to probability theory and mathematical statistics, we obtain

$$\chi^2 = \sum_{i=1}^{16}\frac{m_i^2}{np_i} - n \sim \chi^2(k-r-1), \tag{36}$$

where k is the number of zones, i.e. k=16; r is the number of estimated parameters μ and σ, i.e. r=2. Then, we can determine whether the residual coefficients obey the Laplacian distribution with expected values and variances. According to statistical hypothesis testing, a significance level α refers to the probability of wrongly rejecting the null hypothesis that the distribution is Laplacian. Its corresponding test critical value $\chi^2_\alpha(13)$ can be obtained by checking Laplacian distribution table. The decision of Laplacian distribution is then expressed as

$$\sum_{i=1}^{16}\frac{m_i^2}{np_i} - n < \chi^2_\alpha(13). \tag{37}$$

If the condition defined in Eq. (37) is satisfied, residual coefficients can be assumed to obey a Laplacian distribution. In that case, the depth is considered as optimal, and consequently the other depths are skipped.

In order to significantly improve the coding speed and maintain the coding efficiency, obtaining an optimal value of α and the corresponding $\chi^2_\alpha(13)$ is key. Similarly to mode prediction, we also evenly divide depths into three classes with high, medium and low probabilities. We should select the best threshold value based on these probabilities. Toward this end, some commonly used α values are selected for testing under different depth probabilities. The corresponding coding efficiency represented by BDBR for depths with high probabilities is shown in Table VIII.

From Table VIII, we observe that there is a turning point in sequence "flowervase" when $\chi^2_\alpha(13)$ is equal to 3.565. If $\chi^2_\alpha(13)$ is smaller than or equal to 3.565, the corresponding absolute BDBR in all these sequences are 0.1%. Obviously, the larger $\chi^2_\alpha(13)$ is, the more the encoding speed increases. Based on the above analysis, $\chi^2_\alpha(13)$ is set to 3.565. In the same way, for depths with medium probabilities, the best $\chi^2_\alpha(13)$ is set to 1.783; for depths with low probabilities, the best $\chi^2_\alpha(13)$ is set to 0.446.

Through the above process, the best threshold values $\chi^2_\alpha(13)$ for depths with high, medium and low probabilities are 3.565, 1.783 and 0.446, respectively.

IV. EXPERIMENTAL RESULTS

In order to test the performance of the proposed fast intra prediction algorithm for QS, we have used the SHVC reference

TABLE VIII: Threshold values and the corresponding BDBR

BDBR / threshold values / Sequence	0.892	1.783	3.565	4.107	5.009	5.892
Sunflower	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%
Tractor	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%
Flowervase	-0.10%	-0.10%	-0.10%	-0.20%	-0.40%	-0.50%
PartyScene	0.00%	0.00%	0.00%	-0.10%	-0.10%	-0.10%
BlowingBubble	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RaceHorses	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Parkrunner	0.00%	0.00%	0.00%	-0.10%	-0.10%	0.00%
Town	0.00%	-0.10%	-0.10%	-0.20%	-0.20%	-0.20%

TABLE IX: Performance comparison with QP (22, 26, 30, 34)

Sequence	CBDP		MPPZBB-IMD		GHCB-IMP		DPRDB-DET	
	BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	0.80%	47.69%	-0.3%	81.55%	-0.2%	33.19%	-0.3%	23.65%
PeopleOnStreet	0.60%	47.93%	-0.2%	82.14%	-0.1%	33.87%	-0.2%	21.98%
Kimono	0.20%	56.18%	-0.2%	73.58%	-0.3%	32.58%	-0.1%	22.51%
ParkScene	0.80%	49.40%	-0.2%	69.07%	-0.1%	32.86%	-0.3%	26.52%
Cactus	0.50%	49.02%	-0.2%	70.74%	-0.2%	32.59%	-0.1%	22.34%
BasketballDrive	0.30%	47.26%	-0.3%	69.29%	-0.1%	32.99%	-0.1%	20.17%
BQTerrace	0.70%	52.65%	-0.3%	68.93%	-0.1%	33.70%	0.0%	21.03%
Average	0.56%	50.02%	-0.26%	73.61%	-0.16%	33.11%	-0.16%	22.6%

software (SHM 11.0) on a server with Intel (R) 2.0 GHz CPU and 30 GB memory for testing. The experimental parameters are set according to the CSTC [39]. Since the proposed algorithm is developed for all intra coding, both GOPSize and IntraPeriod are set to be 1. The QPs in the BL are (26, 30, 34, 38), and the corresponding QPs used for the EL are (22, 26, 30, 34) and (20, 24, 28, 32), respectively. To fairly demonstrate the performance of our proposed algorithm, we do not use the eight training sequences in our performance testing. Therefore, our proposed algorithm is generic and effective for different types of video sequences. Coding efficiency is measured by BDBR. A negative BDBR represents an increase in coding efficiency compared with the reference software. Coding speed is evaluated by encoding run-time only in the EL, and "TS" represents the percentage of encoding run-time savings in the EL only.

The proposed algorithm includes four strategies, namely "CBDP", "MPPZBB-IMD", "GHCB-IMP", and "DPRDB-DET". Our experimental results show that the EL with (22, 26, 30, 34) and (20, 24, 28, 32) provide similar performances, therefore only the results for the EL with (22, 26, 30, 34) are provided in this section. The performances of the different strategies are shown in Table IX.

From Table IX, the average coding speed improvements in "CBDP", "MPPZBB-IMD", "GHCB-IMP" and "DPRDB-DET" are 50.02%, 73.61%, 33.11%, and 22.6%, respectively. In parallel, the average BDBR in "CBDP", "MPPZBB-IMD", "GHCB-IMP" and "DPRDB-DET" are 0.56%, -0.26%, -0.16%, and -0.16%, respectively. Since every depth includes the whole intra prediction and ILR prediction process, "CBDP" reaches remarkable coding speed improvements. As the coding process of intra is very complex, skipping intra coding by checking "MPPZBB-IMD" can also achieve significant computational complexity gains. Since only partial intra modes are checked in both RMD and RDO process, "GHCB-IMP" provides smaller, but nonetheless important, time savings. Since threshold values are set to be very small, DPRDB-DET leads to the least significant time savings.

TABLE X: Performance comparison with Q1

Sequence	Proposed		EMSIP [38]		PAPS [35]	
	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	-0.4%	86.03%	-0.2%	72.35%	1.28%	72.02%
PeopleOnStreet	-0.2%	85.49%	-0.1%	70.80%	1.91%	60.36%
Kimono	-0.1%	84.27%	-0.2%	77.06%	0.78%	51.96%
ParkScene	-0.2%	80.79%	0.0%	75.87%	0.65%	63.78%
Cactus	-0.2%	81.16%	-0.1%	76.48%	0.98%	58.42%
BasketballDrive	-0.2%	81.67%	0.2%	81.87%	1.33%	59.36%
BQTerrace	0.0%	81.40%	-0.1%	72.88%	0.65%	58.23%
Average	-0.19%	82.97%	-0.07%	75.33%	1.16%	60.59%

We have compared the overall performance of our algorithm, including the integration of these four strategies, with PAPS algorithm [35] and EMSIP algorithm [38]. For fair comparisons, we have tested all algorithms on the same computing platform. Two settings of QPs: Q1 = (22, 26, 30, 34) and Q2 = (20, 24, 28, 32) are used. The overall performance comparisons in terms of the coding efficiency and the coding speed are shown in Table X and Table XI, with Q1 and Q2, respectively.

From Table X, we find that the averages of BDBR and TS of the proposed algorithm are -0.19% and 82.97% respectively, those of the EMSIP algorithm are -0.07% and 75.33%, and those of the PAPS algorithm are 1.16% and 60.59%. From Table XI, we observe that the averages of BDBR and TS of the proposed algorithm are -0.17% and 81.59% respectively, those of the EMSIP algorithm are -0.17% and 75.07%, and

TABLE XI: Performance comparison with Q2

Sequence	Proposed		EMSIP [38]		PAPS [35]	
	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	-0.4%	84.23%	-0.3%	71.63%	1.30%	71.56%
PeopleOnStreet	-0.3%	84.02%	-0.3%	71.08%	1.94%	68.78%
Kimono	-0.2%	83.57%	0.0%	77.14%	0.80%	52.23%
ParkScene	-0.2%	80.04%	0.0%	76.08%	1.20%	63.78%
Cactus	-0.2%	79.50%	-0.2%	76.51%	0.96%	58.34%
BasketballDrive	0.1%	80.35%	-0.2%	81.40%	1.32%	59.64%
BQTerrace	0.0%	79.44%	-0.2%	71.63%	0.67%	59.13%
Average	-0.17%	81.59%	-0.17%	75.07%	1.17%	60.79%

those of the PAPS algorithm are 1.17% and 60.79%. Compared with EMSIP algorithm, the BDBR of the proposed algorithm is equal or better. Meanwhile, our algorithm is faster than EMSIP. Compared with PAPS algorithm, the proposed algorithm outperforms PAPS in terms of BDBR. Simultaneously, our algorithm is significantly faster than PAPS. Therefore, the proposed algorithm significantly improves the performance, both in coding speed and coding efficiency, when compared to EMSIP and PAPS algorithms.

The main reasons why the proposed algorithm can significantly improve the coding speed are: (1) Since the current CU in EL and the co-located CU in BL are the same except their QPs in quality SHVC, the inter-layer correlation is very strong and ILR mode can predict the corresponding CU in EL very well. Therefore, many CUs only use smaller depths in prediction. We develop CBDP to efficiently predict smaller depths and skip many larger ones. Since the complexity of coding larger depths is very high, skipping these depths can significantly improve the coding speed; (2) Since the ILR correlation is very strong, most CUs use ILR mode in prediction. In addition, the coding process of ILR mode is very simple and takes less time. We develop MPPZBB-IMD to predict ILR mode and skip time-consuming intra encoding process which includes RMD and RDO; (3) Since intra mode represents only a small proportion, we develop GHCB-IMP to only encode a small part of intra mode in RMD, and early terminate the RDO process. Therefore, much coding time can be saved. In addition, a small portion of intra mode leads to a very small coding efficiency lose; (4) We use DPRDB-DET to early terminate depth selection to further improve the coding speed.

In short, the main reason for obtaining our experimental results is that many CUs use smaller depths and most CUs adopt ILR mode, both of which are very simple and take little time. As we develop the proposed algorithm to predict these smaller depths and ILR mode, many larger depths and intra modes can be skipped accordingly.

From Table X and Table XI, we can find that the BDBR of our scheme can achieve BDBR savings of 0.19% and 0.17% in average, respectively, compared against the SHM reference software. Generally speaking, the improvements of coding speed usually lead to BDBR increase, namely coding efficiency decrease. But it is possible to also obtain BDBR decrease, i.e., coding efficiency increase. One major reason is the strong dependency among CUs, mainly due to the extensive use of predictive coding. For the proposed algorithm, the explanation lays in the intra prediction process. The results of our investigation and analysis are presented below.

TABLE XII: RD costs in MPPZBB-IMD and SHM reference software

MPPZBB-IMD		SHM	
521(UL)	3398(U)	521(UL)	3392(U)
1311(L)	4321(C)	1306(L)	4405(C)

TABLE XIII: RD costs and modes of different methods

Number	SHM		MPPZBB-IMD		Intra-only	ILR-only
	RD cost	mode	RD cost	mode	RD cost	RD cost
1	362	intra	360	intra	324	392
2	641	intra	628	intra	692	717
3	272	intra	277	intra	250	286
4	326	ILR	318	intra	344	326

To investigate the RD costs, we select the MPPZBB-IMD to test the sequence "Flowervase", with QP=22 in EL. The RD costs in PZBB-IMD and the SHM reference software for the 4 CTUs (2 rows and 2 columns) with the upper left location at (64, 64) in the first frame are shown in Table XII. As shown in Table XII, three RD costs of our method are equal or slight larger than the corresponding RD costs of SHM. However, the RD cost of the bottom right CTU by our method is significantly smaller than its counterpart in the SHM. The sum of the 4 RD costs in our method and SHM are 9552 and 9616, respectively. We have also calculated the sum for the whole first frame. The sum with our method is also smaller than that with SHM. Thus, our scheme can occasionally achieve BDBR decrease when compared to SHM.

Since the larger CUs may include different modes, only 8×8 CUs are selected in our investigation. We further use intra-only mode and ILR-only mode in evaluating RD costs and modes. Table XIII lists the corresponding RD costs and modes of four typical 8×8 CUs.

From Table XIII, we can observe that both SHM and MPPZBB-IMD use intra mode in the first three CUs, but their corresponding RD costs are different from those of intra-only method in each CU. In the first two CUs, RD costs in MPPZBB-IMD are smaller than those in SHM. In the third CU, RD cost in MPPZBB-IMD is larger than that in SHM. From the first three CUs, we can observe that even the same CU uses intra mode in prediction, its corresponding RD costs are different. In the fourth CU, SHM uses ILR mode, MPPZBB-IM uses intra mode, and RD costs in MPPZBB-IM and intra-only are different. Therefore, we can conclude that not only the RD cost changes, but the mode may also change in different methods.

The reason is that CUs are predicted by their reference pixels in intra prediction, as shown in Fig.5. When reference pixels are more similar with the texture of CUs, intra mode can predict more accurately and the corresponding RD costs are smaller, and vice versa. In Fig. 5, using different methods may lead to different predictions in CU L and U, so the reference pixels L-pixels and U-pixels, shown in black dots, will also be different. Obviously, it will lead to different RD costs for

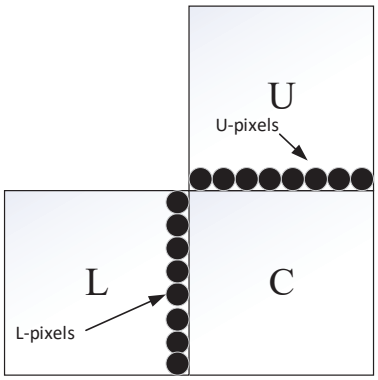

Fig. 5: Intra prediction in CTU

CU C. If L-pixels and U-pixels in SHM are more similar with the texture of CU C, SHM can predict more accurately and the corresponding RD costs are smaller, such as the first two CUs in Table XIII. Otherwise, MPPZBB-IM can predict more accurately and the corresponding RD costs are smaller, e.g., the third CU in Table XIII. Similarly, in Table XII, RD Costs of CTU L and CTU U with SHM are smaller than those with MPPZBB-IMD, thus SHM can provide a better prediction for CTU L and CTU U than MPPZBB-IMD. However, left reference pixels and upper pixels in SHM cannot with certainty predict CTU C better than those in MPPZBB-IMD method. Thus, MPPZBB-IMD method can still occasionally achieve a little BDBR decrease.

In summary, in intra prediction, different methods may lead to differences in reference pixels, and the corresponding RD costs are also different. Compared to SHM, our method may therefore occasionally achieve either a decrease or an increase in BDBR.

V. Conclusion

In this paper, we have proposed a new and effective intra prediction algorithm for Quality SHVC. The encoding procedure includes four fast strategies: "CBDP" combines spatial and temporal correlations as well as their correlation degrees in Naive Bayes classifier to obtain depth probabilities, and skip depths with low likelihood; "MPPZBB-IMD" combines with ILR mode probabilities and SSD based PZBs to skip unnecessary intra prediction; "GHCB-IMP" uses Sobel operator to theoretically obtain two gradients with greatest amplitudes as initial modes to search the first N candidate modes in RMD, which are early terminated based on the HC values of two neighboring IMs in RDO. Thus, only a subset of IMs instead of all 35 IMs are used in both RMD and RDO processes; and "DPRDB-DET" jointly uses depth probabilities and residual distribution for early termination of depth selection. Therefore, the proposed algorithm can significantly accelerate the coding speed with negligible decreases in coding efficiency. In our future research activities, we will use deep learning to further improve the coding speed of SHVC.

References

[1] L. Q. Shen, and Z. Y. Zhang,"Content-Adaptive Motion Estimation Algorithm for Coarse-Grain SVC," *IEEE Trans. Image Process.*, vol.21, no.5, pp.2582-2591, May. 2012.

[2] G. Correa, P. Assuncao, L. Agostini, and L. S. Cruz, "Performance and Computational Complexity Assessment of High-Efficiency Video Encoders,"*IEEE Trans. Circuits Syst. Video Technol.*, vol.22, no.12, pp.1899-1909, Dec. 2012.

[3] J. M. Boyce, Y. Ye, and J. L. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol.26, no.1, pp.20-34, Jan. 2016.

[4] H. Li, Z. G. Li, C. Y. Wen, and L. P. Chau, "Fast mode decision for spatial scalable video coding," *in Proc. IEEE Int. Symp. Circuits Syst.*, Kos, Greece, May 2006, pp. 3005–3008.

[5] H. Li, Z. G. Li, and C. Wen, "Fast mode decision algorithm for interframe coding in fully scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 7, pp. 889–895, Jul. 2006.

[6] H. C. Lin, W. H. Peng, and H. M. Hang, "A fast mode decision algorithm with macroblock-adaptive rate-distortion estimation for intra only scalable video coding," *in Proc. IEEE Int. Conf. Multimedia Expo.*, Hannover, Germany, Jun. 2008, pp. 765–768.

[7] Bumshik Lee, and Munchurl Kim. "A Low Complexity Mode Decision Method for Spatial Scalability Coding," *IEEE Trans. Circuits Syst. Video Technol.*, 2011,21(1): 88-95.

[8] Byung-Gyu Kim, Krishna Reddy, Yoon-Young Park. Fast mode decision algorithm for inter-frame coding in H.264 extended Scalable Video Coding. *IEEE International Symposium on Circuits and Systems.*, 2009,5:2345-2348.

[9] L. Shen and Z. Zhang, "Content adaptive motion estimation algorithm for coarse-grain scalability," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 2582–2591, Mar. 2012

[10] L. Shen, Y. Sun, Z. Liu, and Z. Zhang, "Efficient skip mode detection for coarse grain quality scalable video coding," *IEEE Signal Process. Lett.*, vol. 17, no.10, pp.887–890, Oct. 2010..

[11] C. S. Park, B. K. Dan, H. Choi, and S. J. Ko, "A statistical approach or fast mode decision in scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 12, pp. 1915–1920, Dec. 2009.

[12] C. H. Yeh, K.J. Fan, M.J. Chen, and G.-L. Li, "Fast mode decision algorithm for scalable video coding using Bayesian theorem detection and Markov process," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 4, pp. 563–574, Apr. 2010.

[13] S. W. Jung, S.J.Baek, C.S.Park, and S.-J. Ko, "Fast mode decision using all-zero block detection for fidelity and spatial scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 201–206, Feb. 2010.

[14] T. S. Zhao, S. Kwong, H. L. Wang, and C. C. Kuo, "H.264/SVC mode decision based on optimal stopping theory," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2607–2618, May 2012..

[15] X. Lu and G. R. Martin, "Fast mode decision algorithm for the H.264/AVC scalable video coding extension," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 1, pp. 846–855, Jan. 2013.

[16] D. Y. Wang, C. Yuan, Y. Sun, J. Zhang, and X. Jin, "A fast mode decision algorithm applied to coarse-grain quality scalable video coding," *J. Vis. Commun. Image Represent.*, vol. 25, no. 7, pp. 1631-1639, Jul. 2014.

[17] W. Jiang, H. Ma, and Y. Chen, "Gradient based fast mode decision algorithm for intra prediction in HEVC," *in Proc.2nd Int. Conf. Consumer. Electron.*, Commun. Netw., Apr.2012, pp.1836-1840.

[18] S. Q. Yan, L. Hong, W. F. He and Q.Wang, "Group-Based Fast Mode Decision Algorithm for Intra Prediction in HEVC," *IEEE International Conference on Signal Image Technology and Internet Based Systems.*, Naples, Italy, November. 25-29, 2012.

[19] H. Zhang, and Z. Ma, "Fast Intra Mode Decision for High-Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol.24, no.4, pp.660-668, Apr.2014.

[20] B. Min, and R.C.C. Cheung, "A Fast CU Size Decision Algorithm for HEVC Intra Encoder," *IEEE Trans. Circuits Syst. Video Technol.*, vol.25, no.5, pp.892-896, May. 2015.

[21] S. Cho, and M. Kim, "Fast CU Splitting and Pruning for Suboptimal CU Partitioning in HEVC Intra Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol.23, no.9, pp.1555-1564, Sep. 2013.

[22] L. Q. Shen, Z. Y. Zhang, and P. An, "Fast CU Size Decision and Mode Decision Algorithm for HEVC Intra Coding," *IEEE Trans.Consumer Electronics.*, vol.59, no.1, pp.207-213, Jan. 2013.

[23] M. Jamali, S. Coulombe, F. Caron, “Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification,” *Proceedings of Data Compression Conference*, Snowbird, UT, USA, July 2015.

[24] K. Lim, J. Lee, S. Kim, and S. Lee, “Fast PU skip and split termination algorithm for HEVC intra prediction,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1335–1346, Aug. 2015.

[25] N. Hu and E. H. Yang, “Fast mode selection for HEVC intra-frame coding with entropy coding refinement based on a transparent composite model,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1521–1532, Sep. 2015..

[26] J. Chen and L. Yu, “Effective HEVC intra coding unit size decision based on online progressive Bayesian classification,” *in Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Seattle, WA, USA, Jul. 2016, pp. 1–6..

[27] B. Du, W. C. Siu, and X. Yang, “Fast CU partition strategy for HEVC intra-frame coding using learning approach via random forests,” *in Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC).*, Hong Kong, Dec. 2015, pp. 1085–1090..

[28] D. R. Coll, V. Adzic, G. F. Escribano, H. Kalva, J. L. Martínez, and P. Cuenca, “Fast partitioning algorithm for HEVC intra frame coding using machine learning,” *in Proc. IEEE Int. Conf. Image Process.(ICIP)*, Paris, France, Oct. 2014, pp. 4112–4116.

[29] T. Zhang, M. T. Sun, D. Zhao, and W. Gao, “Fast intra-mode and CU size decision for HEVC,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 8, pp. 1714–1726, Aug. 2016.

[30] 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 Trans. Image Process.*, vol. 25, no. 11, pp. 5088–5103, Nov. 2016.

[31] H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, “Content adaptive complexity reduction scheme for quality/fidelity scalable HEVC,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP).*, Vancouver, BC, Canada, May 2013, pp. 1744–1748

[32] H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, "Content Adaptive Complexity Reduction Scheme for Quality/Fidelity Scalable HEVC," *document JCTVC-L0042, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11.*, Geneva, Switzerland, Jan. 2013.

[33] H. R. Tohidypour, H. Bashashati, M. T. Pourazad, and P. Nasiopoulos, “Fast mode assignment for quality scalable extension of the High Efficiency Video Coding (HEVC) standard: A Bayesian approach,” *in Proc. 6th Balkan Conf. Inf. (BCI).*, Thessaloniki, Greece, Sep. 2013, pp. 61–65

[34] H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, “An encoder complexity reduction scheme for quality/fidelity scalable HEVC,” *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 664–674, Sep. 2016.

[35] H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, “Probabilistic approach for predicting the size of coding units in the quad-tree structure of the quality and spatial scalable HEVC,” *IEEE Trans. Multimedia.*, vol. 18, no. 2, pp. 182–195, Feb. 2016..

[36] H. R. Tohidypour, H. Bashashati, M. T. Pourazad, and P. Nasiopoulos,"Online-Learning-Based Mode Prediction Method for Quality Scalable Extension of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no.10, pp.2204-2215, Oct. 2017..

[37] D. Wang, C. Yuan, Y. Sun, J. Zhang, and H. Zhou, "Fast mode and depth decision algorithm for intra prediction of quality SHVC," *Intell. Comput. Theory, ser. Lecture Notes in Comput. Sci.*, vol.8588, pp.693–699, 2014.

[38] D. Wang, C. Zhu, Y. Sun, F. Dufaux, Y. Huang, "Efficient Multi-Strategy Intra Prediction for Quality Scalable High Efficiency Video Coding," *IEEE Trans. Image Process.*, vol.28, no.4, pp.2063-2074, Apr. 2019.

[39] Common SHM Test Conditions and Software Reference Configurations, *Doc. JCTVC-Q1009, ITU-T SG 16 WP 3 and ISO/IEC JTC1/SC 29/WG.*, 11, Mar. 2014.

[40] D. Wang, Y. Sun, W. Li, C. Zhu, F. Dufaux, “Fast Inter Mode Predictions for SHVC,” *in Proc. IEEE Int. Conf. Multimedia & Expo(ICME)*, Shanghai, China, Jul. 2019, pp. 1696-1701.

[41] D. Wang, Y. Sun, W. Li, C. Zhu, F. Dufaux, "Fast Depth and Inter Mode Prediction for Quality Scalable High Efficiency Video Coding," *IEEE Trans. multimedia.*(In Press, DOI:10.1109/TMM.2019.2937240).

[42] G.Bjontegaard, "Calculation of average PSNR difference between RD-curves," *13th VCEG-M33 Meeting.*, Austin, TX, Apr.2-4, 2001.

[43] I. Pao, M. Sun, "Modeling DCT coefficient for fast video encoding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp.608- 616, Apr.1999