



Probability-Based Fast Intra Prediction Algorithm for Spatial SHVC

Dayong Wang, Yu Sun, Jinhua Liu, Frédéric Dufaux, Xin Lu, Bo Hang

► To cite this version:

Dayong Wang, Yu Sun, Jinhua Liu, Frédéric Dufaux, Xin Lu, et al.. Probability-Based Fast Intra Prediction Algorithm for Spatial SHVC. IEEE Transactions on Broadcasting, 2022, 68 (1), pp.83 - 96. 10.1109/TBC.2021.3126277 . hal-03432049

HAL Id: hal-03432049

<https://hal.science/hal-03432049>

Submitted on 17 Nov 2021

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.

Probability-Based Fast Intra Prediction Algorithm for Spatial SHVC

Dayong Wang, Yu Sun, Jinhua Liu, Frederic Dufaux, *Fellow, IEEE*, Xin Lu, and Bo Hang

Abstract—Due to multi-layer encoding and Inter-layer prediction, Scalable High Efficiency Video Coding (SHVC) has extremely high coding complexity. It is very crucial to improve its coding speed so as to promote widespread and cost-effective SHVC applications. In this paper, we propose a new probability-based fast Intra prediction algorithm for spatial SHVC. More specifically, first, we integrate depth probabilities with textural based all-zero blocks and all-nonzero blocks through Lagrange Interpolation Polynomial (LIP) to derive thresholds to early skip unlikely depths and early terminate depth selection. Second, to early skip Intra mode prediction, we combine mode probabilities with Jarque-Bera test through LIP to test the residual coefficients of Inter-layer reference (ILR) mode. Third, we adopt the difference of Hadamard Costs (HCs) of horizontal, vertical and diagonal directional modes to predict candidate directional modes (DMs), and then exploit both the percentages of gradient amplitudes and HCs for DM early termination. Experimental results demonstrate that the proposed algorithm can achieve a speed up gain of more than 61% with 0.03% decrease in Bjøntegaard Delta Bit Rate on average.

Index Terms—SHVC, depth decision, ILR mode, directional mode, early termination.

I. INTRODUCTION

VIDEO applications, such as digital TV broadcasting, video conferencing and smartphone communications, are so prevalent in our life right now, especially during the COVID-19 Pandemic Period. On one hand, diverse devices

This work was supported in part by Natural Science Foundation of Chongqing under Grant cstc2020jcyj-msxmX0766; in part by the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant KJZD-K202100604; in part by the Jiangxi Provincial Natural Science Foundation under Grant 20202BABL202006 and 20192BAB207013; in part by International Science and Cooperation Project of Hubei Province under Grant 2019AHB059; in part by Science and Technology Development Program of Central Guide to Local Government of China under Grant 2019ZYYD043; and in part by faculty sabbatical leave fund from University of Central Arkansas. (*Corresponding author: Jinhua Liu*)

D. Wang is with the Chongqing Key Laboratory on Big Data for Bio Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China, and also with the School of mathematics and computer science, Shangrao Normal University, Shangrao 334001, China (e-mail: wang-dayong@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).

J. Liu is with the School of Mathematical and Computer Sciences, Shangrao Normal University, Shangrao 334001, China. (e-mail: liujinhua_uestc@126.com).

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).

X. Lu is with the Faculty of Computing, Engineering and Media (CEM), De Montfort University, Leicester, LE1 9BH, United Kingdom (e-mail: xin.lu@dmu.ac.uk).

B. Hang is with the Computer school, Hubei university of arts and science, Xiang Yang 441053, Hubei, China (e-mail: bohong@163.com).

Manuscript received Aug 2, 2021

lead to various screen resolutions, processing capabilities, and network bandwidth requirements. On the other hand, different networks, such as Broadband networks, WIFI networks, and 4G/5G wireless networks, may have time-varying bandwidths. Consequently, video streaming must have a very strong adaptation ability. Scalable High Efficiency Video Coding (SHVC) is an efficient solution to address this requirement. SHVC supports temporal scalability, spatial scalability, quality scalability, bit-depth scalability and color gamut scalability. It consists of a base layer (BL) and one or more Enhancement layers (ELs). Through selecting an appropriate EL, SHVC can adapt to various devices and network conditions.

SHVC is a scalable extension of High Efficiency Video Coding (HEVC). Although HEVC has very high coding efficiency, it also has very high encoding complexity. Compared to H.264/AVC, the encoding complexity of HEVC is increased by about two to four times [1]. SHVC needs to encode a BL and one or more ELs. In addition to encoding Inter-layer prediction, each EL also needs to be further encoded using the same encoding process as HEVC. Thus, SHVC is much more complex than HEVC. Due to the high encoding complexity, the wide applications of SHVC have been limited, especially for wireless and real-time applications. Therefore, it is very important to improve coding speed.

For this purpose, in this paper, we propose a new probability-based fast Intra prediction algorithm for spatial SHVC. First, we obtain temporal and spatial correlation degrees between two successive frames, and use a Bayesian formulation to obtain depth probabilities, which are integrated with textural based all-zero blocks and all-nonzero blocks (AZBANBs) through Lagrange Interpolation Polynomial (LIP) to early skip unlikely depths and early terminate depth selection. Second, using the similar method of obtaining depth probabilities, we derive mode probabilities, and combine them with Jarque-Bera test through LIP to early skip Intra mode prediction. Third, in the rough mode decision (RMD) procedure, we adopt the difference of Hadamard Costs (HCs) of horizontal, vertical and diagonal directional modes to predict candidate directional modes (DMs), and then simultaneously exploit the percentages of gradient amplitudes and HCs for early termination.

The major novelties and contributions of the proposed algorithm are summarized below:

- (1) For different scalability ratios, 1.5x and 2x, in spatial SHVC, we develop a new method to obtain temporal and spatial correlation degrees between two successive frames.
- (2) By combining textural complexity with quantization

parameters (QPs), we develop an approach to derive a textural based AZBANBs for depth selection.

- (3) We propose to adopt LIP to derive thresholds in prediction on early skip and termination for depths, Inter-layer reference (ILR) mode and DMs.
- (4) We adopt the differences of HCs of horizontal, vertical and diagonal DMs to predict candidate DMs, and develop an early termination method for DMs based on the percentages of gradient amplitudes and HCs.

The rest of this paper is organized as follows. Section II discusses related work. Section III provides the overview of the proposed approach. Section IV presents three fast strategies to speed up encoding. Section V discusses and analyses the experimental results. Finally, Section VI draws the conclusions of this research and plans for future work.

II. RELATED WORK

Due to the fact that SHVC is a scalable extension of HEVC, there should be a strong relationship between SHVC and HEVC. Therefore, in this section, we first briefly discuss fast coding algorithms for HEVC, and then review the algorithms to speedup SHVC.

Some fast algorithms focused on improving the coding speed of HEVC have been developed in [2]-[6]. Textural features and correlations have relationships with mode and depth selections. The work in [2] studies textural features to predict candidate modes and depths. The work in [3] adopts temporal and spatial correlations to predict candidate depths and skip unlikely depths. The research in [4] splits residual coefficients of a coding unit (CU) into top and bottom parts as well as left and right parts, then decides whether to further split this CU according to the difference of two parts of each division. The work in [5] first obtains the difference between original luminance pixels and predicted ones, and then develop two decision models to early skip of vertical binary-tree and horizontal and vertical ternary-tree partition for each CU. The work in [6] first develops an RDO cost statistical model, and use HCs to predict candidate DMs and skip unlikely DMs based on the model.

As we know, unlike HEVC, SHVC also has Inter-layer correlation. Therefore, fast coding algorithms for HEVC can not be applied directly for SHVC coding speedup. In other words, algorithms specifically designed for HEVC, and used straightforwardly in coding speedup for SHVC, cannot reach the optimal performance. Therefore, it is crucial to design fast coding algorithms for SHVC based on its own coding structure.

A number of fast coding algorithms are developed for SHVC [7]-[22]. Tohidypour et al. [7] use relative CUs' Rate Distortion (RD) to predict the current CU's RD cost in EL for early termination. To reduce the coding complexity, the technique proposed in [8] uses relative CUs to predict likelihood modes and skip unlikely modes in EL. According to the combination of depth and mode of the co-located CU in BL, the algorithm proposed in [9] first predicts likelihood modes and excludes unlikely modes of the current CU in EL. Then, it further eliminates unlikely modes based on Inter-layer and spatial correlations. In [10], according to spatial

and Inter-layer correlations, a Naïve Bayesian Classifier is adopted to predict the quad-tree structure of coding tree units (CTUs) for SHVC. Based on the coding information of relative CUs, a method about online-learning-based mode prediction is developed in [11] to predict the likelihood modes of the current CU in EL. Lu et al. [12] use Inter-layer and spatial correlations as well as textural complexity to predict candidate coding depths. Lu et al. [13][14] jointly use texture complexity and spatio-temporal correlation to predict candidate depths. Then, they combine Inter-layer correlation with temporal correlation to exclude unlikely DMs. Exploiting correlations in prediction is one common attribute of the above algorithms.

However, different CUs may have different correlation degrees. Using correlations only, without considering correlation degrees, can certainly influence the performance. To improve the coding speed, Wang et al. [15, 16, 17] exploit both correlations and correlation degrees to predict candidate depths. Then, they use residual coefficients and RD costs to early terminate mode and depth selection. Different depths and modes may have different probabilities to be selected. However, these works do not consider probabilities to early terminate modes or depth selection, which limit the improvement of coding speed. In order to address this issue, the work in [18] predicts the probabilities of depths and modes based on correlations and correlation degrees, and then evenly divides the probabilities into three groups. In each group, a fixed threshold is used to early terminate ILR mode and depth selection. The above works are developed for quality SHVC. Since the resolutions in BL and EL are the same, the correlation degrees in EL can be obtained directly based on co-located correlation degrees in BL. The works in [19, 20] obtain correlation degrees based on pixel correlations. Based on Inter-layer, spatio-temporal and Inter-level correlations, the research [19] develops a conditional probability of a SKIP/Merge mode, motion activity and mode complexity to predict candidate modes. The Inter-layer and spatio-temporal correlation are adopted in [20] to build two feedforwards neural network-based learning models to predict depths and modes. In brief, one common feature about the above works [15-20] is that correlation degrees are fully utilized in prediction.

Clearly, the fact that different depths and modes usually have different probabilities should be considered to obtain more accurate thresholds in prediction. The work in [18] divides probabilities into several groups and set a fixed value for each group in prediction. However, probabilities in each group are also different, which also requires variable thresholds instead of fixed ones. Therefore, to further improve the performance, it is highly desirable to develop probabilities-based algorithm in prediction.

The above techniques are mainly developed for quality SHVC, except several works [10, 12-14] that are applied to Intra coding for spatial SHVC. One possible reason is that the resolutions in BL and EL are the same in quality SHVC, whereas they are different in spatial SHVC. It is relatively easier to develop fast coding algorithms for quality SHVC due to a stronger Inter-layer correlation. Undoubtedly, how to develop effective and fast coding algorithms for spatial SHVC is definitely a research challenge, due to the following

considerations:

- (1) In order to predict more accurately, we need to obtain temporal and spatial correlation degrees, which can be generally obtained through two ways. One way is to obtain correlation degrees through pixel correlations [19][20]. Obviously, depth and mode selections are not only correlated with pixels, but are also correlated with reference pixels and QPs. Therefore, only using pixels in calculating correlation degrees may not always obtain the optimal performance. The other way is to use co-located correlation degrees in BL for quality SHVC [15-18]. However, in spatial SHVC, these correlation degrees cannot be obtained directly based on co-located correlation degrees in BL, since the resolutions in BL and EL vary between two scalability ratios (1.5x and 2x). Thus, new methods to obtain temporal and spatial correlation degrees for spatial SHVC must be explored.
- (2) Both textural complexity and QPs have strong relationships with depth selection. However, some works only consider textural complexity, such as [12-14]. Many works establish relation among textural complexity, QP and depth selection based on experiments [21]. Due to the diversity of video sequences, these empirical relations may not be optimum for all sequences. Apparently, a threshold obtained theoretically will have strong adaptive ability for various videos. Therefore, we propose texture based AZBANBs to derive the threshold for depth selection.
- (3) Different depths and modes may lead to different probabilities. Many research works have not considered depth and mode probabilities in prediction [15-17]. Thus, they certainly cannot obtain the optimum performance. The work [18] divides probabilities on depth and ILR into three groups. Each group has its fixed threshold for depth and ILR mode early termination. This method obtains better performance than the ones without considering probabilities. However, probabilities in the same group may be different, their corresponding thresholds should also be different. Apparently, using the same fixed threshold for different probabilities in the same group cannot achieve the optimum performance. Therefore, an approach on deriving variable thresholds for different probabilities should be developed so as to further improve the performance in prediction.
- (4) In order to obtain candidate DMs in RMD procedure, many works use textural features in prediction [21]. The selection of DMs is not only related to textural features, but also related to reference pixels. Obviously, only using textural features cannot predict DMs very well. HCs are used in [15, 18] to predict candidate DMs. However, only considering HC values without considering HC differences will definitely limit the improvement of coding speed.

In order to address the above issues, we propose the corresponding solutions and develop a new probability-based fast Intra prediction algorithm for spatial SHVC. According to the best of our knowledge, the textural based AZBANBs and LIP have never been applied in video coding speedup.

Therefore, the main novelty of this work is to apply them for improving coding speed of SHVC.

III. OVERVIEW OF THE PROPOSED ALGORITHM

In order to improve Intra coding speed and maintain coding efficiency for spatial SHVC, we propose three strategies: Probability and Textural Complexity-Based Depth Prediction (PTC-BDP), Probability and Residual Coefficients-Based ILR Mode Prediction (PRC-BIMP), Probability and HCs-Based DM Prediction (PH-BDP). First, we predict depth candidates through PTC-BDP. For depth candidates, to skip Intra prediction, we determine whether the ILR mode is the best mode though PRC-BIMP. Otherwise, we predict DMs candidates in the RMD process and Rate Distortion Optimization (RDO) process through PH-BDP.

As shown in Fig.1, the left side shows the three strategies and the right side illustrates the procedure of the proposed algorithm.

In PTC-BDP, we obtain temporal and spatial correlation degrees between two successive frames, and combine relative CUs' information with correlation degrees through a Naive Bayesian classifier to derive depth probabilities. The latter are then integrated with texture based AZBANBs through LIP to early skip unlikely depths and early terminate depth selection. In PRC-BIMP, using the similar way as depth prediction, we can obtain mode probabilities, which are combined with Jarque-Bera test through LIP to skip Intra prediction. In PH-BDP, we predict candidate DMs based on the differences of HCs of horizontal, vertical and diagonal DMs. Then, we integrate the percentages of gradient amplitudes with HCs through LIP to early terminate RMD process. Finally, we use the differences of HCs of neighboring DMs to early terminate the RDO process.

IV. THE PROPOSED FAST INTRA PREDICTION PROCESS

In order to develop the above three strategies, we have conducted extensive experiments to investigate the features and characteristics of Intra coding in spatial SHVC. According to common SHM test conditions (CSTC) [23], there exist two scalability ratios, 2x and 1.5x, in spatial SHVC, representing the ratios of both height and width in EL to those in BL. Out of classes A, B, C, D, and E, only test sequences in class B can meet the requirement of two scalability ratios and are thus selected in our experiments. More precisely, the selected sequences are Blue_Sky, Ducks, Park_Joy, Pedestrian, Sunflower, Town, Tractor and Tree. These sequences are representative since they cover motion and texture from simple to complex. Both scalability ratios also include one QP set in BL and two QP sets in EL. The QP set in BL is (22, 26, 30, 34), and the corresponding two QP sets in EL are (22, 26, 30, 34) and (24, 28, 32, 36), respectively. Obviously, the scalability ratio 2x has a larger difference in resolution between BL and EL, and the QP set (24, 28, 32, 36) in EL also has a larger difference in QPs between BL and EL. Therefore, the combination of 2x and the QP set (24, 28, 32, 36) in EL should have the weaker Inter-layer correlation for larger differences

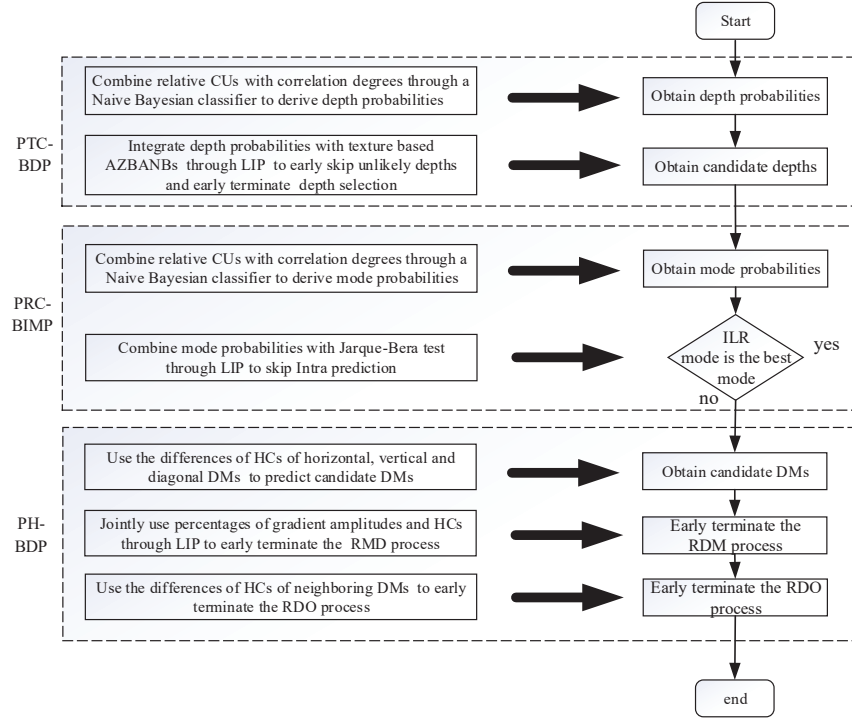


Fig. 1: Overview of the Proposed Algorithm.

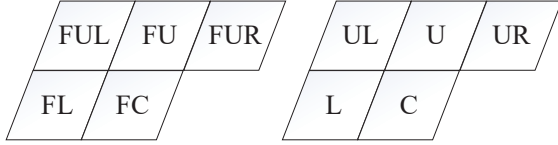


Fig. 2: Relative CUs of the current CU in prediction.

in both scalability ratios and QP sets. If we could achieve good performance in this case, we should reach even better performances in the other configurations. Therefore, we only use 2x and the QP set (24, 28, 32, 36) in EL when conducting experiments. Based on these experiments, we develop three fast Intra prediction strategies which are described below.

A. Probability and Textural Complexity-Based Depth Prediction (PTC-BDP)

In this section, we first obtain depth probabilities based on correlation. Then, we propose texture based AZBANs to measure textural complexities. Finally, we jointly use depth probabilities and texture based AZBANs through LIP to early skip unlikely depths and early terminate depth selection.

1) *Correlation-Based Depth Probability Prediction*: Since the current CU and neighboring CUs are usually very similar, their spatial correlation is very high. Due to the continuity of a video sequence, the temporal correlation of successive frames is also very high. Therefore, we can use spatial and temporal correlations to obtain depth probabilities. The spatial and temporal correlations are exploited through the relative CUs of the current CU, as shown in Fig. 2. The current CU

in the EL, its left CU, upper CU, upper-left CU and upper-right CU are denoted as C, L, U, UL and UR, respectively. Accordingly, the co-located CUs of C, L, U, UL and UR in the previous frame are denoted as FC, FL, FU, FUL and FUR, respectively. We refer to the depth level of C, L, U, UL, UR, FC, FL, FU, FUL and FUR as $d_0, d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9$, respectively.

Obviously, different neighboring CUs may have different correlation degrees with the current CU. It is better to combine neighboring CUs and their correlation degrees in prediction. Thus, we need to obtain the temporal and spatial correlation degrees first. Since neighboring CUs and the current CU are usually very similar, we can calculate the temporal correlation degree of the current CU through its neighboring CUs. Obviously, if one of the neighboring CUs in the current frame and its co-located CU in the previous frame are more similar, the absolute difference of their depths should be smaller and the temporal correlation degree of the current CU should be stronger, and vice versa. In other words, the absolute difference between the depth of the i -th ($1 \leq i \leq 4$) neighboring CU in the current frame and that of its co-located CU in the previous frame, $|d_i - d_{i+5}|$, is inversely related to the temporal correlation degree of the current CU. Since the maximum value of the depth level is 3 and there are four neighboring CUs, the average value of the absolute differences between the depths of four neighboring CUs in the current frame and those of their co-located CUs in the previous frame are used instead of the depth absolute difference of an individual CU. Then, the temporal correlation degree of the current CU, td ,

is calculated by:

$$td = 3 - \frac{\sum_{i=1}^4 |d_i - d_{i+5}|}{4}. \quad (1)$$

As mentioned above, the current and previous frames are usually very similar, thus we can use spatial correlation degrees of CUs in the previous frame as those of the co-located CUs in the current frame. Suppose sd is the spatial correlation degree vector of four neighboring CUs' depths of the CU FC. It is obvious that the absolute difference between the depth of the CU FC and that of its i -th neighboring CU in the previous frame, $|d_{i+5} - d_5|$ with $1 \leq i \leq 4$, is inversely related to their corresponding spatial correlation degree. Since the maximum value of the depth level is 3, the i -th component of the spatial correlation degree vector, sd_i , is calculated by:

$$sd_i = 3 - |d_{i+5} - d_5|. \quad (2)$$

Through the above process, we can obtain the temporal and spatial correlation degrees. Assume \mathbf{nv} is the depth level vector of four neighboring CUs of the current CU, and its i -th component nv_i 's value is d_i . Based on relative CUs and correlation degrees, we use a Bayesian formulation to derive the probability of depth level d_0 , $f_d(d_0)$, as:

$$f_d(d_0) = p(d_0 | ((\mathbf{nv}, \mathbf{sd}), (d_5, td))) = \frac{p(((\mathbf{nv}, \mathbf{sd}), (d_5, td)), d_0)}{p((\mathbf{nv}, \mathbf{sd}), (d_5, td))} \\ = \frac{p(((\mathbf{nv}, \mathbf{sd}), (d_5, td)) | d_0) p(d_0)}{p((\mathbf{nv}, \mathbf{sd}), (d_5, td))}, \quad (3)$$

where the vector $(\mathbf{nv}, \mathbf{sd})$ refers to the depth levels of neighboring CUs of the current CU and their corresponding spatial correlation degrees, $(\mathbf{d}_5, \mathbf{td})$ refers to the depth level of CU FC and its temporal correlation degree, $p((\mathbf{nv}, \mathbf{sd}), (\mathbf{d}_5, \mathbf{td}))$ is the probability of the vector $((\mathbf{nv}, \mathbf{sd}), (\mathbf{d}_5, \mathbf{td}))$, $p(((\mathbf{nv}, \mathbf{sd}), (\mathbf{d}_5, \mathbf{td})) | d_0)$ is the conditional probability of the vector $((\mathbf{nv}, \mathbf{sd}), (\mathbf{d}_5, \mathbf{td})) | d_0$. If we directly calculate Eq. (3), the process is very complex. In order to address this issue, we use a Naive Bayesian classifier to calculate Eq. (3). Since a Naive Bayesian classifier make a conditional independence assumption, we can assume that components in a vector are independent of each other. In addition, temporal correlation and spatial correlation can also be considered to be independent of each other. Based on the above analysis, we can rewrite (3) as:

$$f_d(d_0) = \frac{p(((\mathbf{nv}, \mathbf{sd}), (d_5, td)) | d_0) p(d_0)}{p((\mathbf{nv}, \mathbf{sd}), (d_5, td))} = \frac{p((\mathbf{nv}, \mathbf{sd}) | d_0) p((d_5, td) | d_0) p(d_0)}{p(\mathbf{nv}, \mathbf{sd}) p(d_5, td)} \\ = \frac{p(d_0) p((d_5, td) | d_0) \prod_{i=1}^4 p((nv_i, sd_i) | d_0)}{p(d_5, td) \prod_{i=1}^4 p(nv_i, sd_i)}, \quad (4)$$

where $p(d_0)$, $p(\mathbf{d}_5, \mathbf{td})$, $p((\mathbf{d}_5, \mathbf{td}) | d_0)$, $p(\mathbf{nv}_i, \mathbf{sd}_i)$ and $p((\mathbf{nv}_i, \mathbf{sd}_i) | d_0)$ are the depth probabilities of d_0 , $(\mathbf{d}_5, \mathbf{td})$, $(\mathbf{d}_5, \mathbf{td}) | d_0$, the i -th ($1 \leq i \leq 4$) component in the vector $(\mathbf{nv}, \mathbf{sd})$ and $(\mathbf{nv}, \mathbf{sd}) | d_0$, respectively. According to the experimental conditions mentioned above, we can obtain these five probabilities, and then we can compute $f_d(d_0)$.

Through the above process, the probabilities of all four depth levels can be obtained. As described above, temporal correlation degrees are obtained from its neighboring CUs, and

spatial correlation degrees are obtained from the co-located CUs in the previous frame. In other words, temporal and spatial correlation degrees are indirectly obtained, so these correlation degrees may not always be very accurate. If we directly use these probabilities in prediction, the corresponding coding efficiencies may be obviously degraded in some sequences. In order to address this issue, we further exploit textural complexity and QPs in prediction.

2) *Textural Based All-zero Blocks and All-nonzero Blocks:* Depth selection has strong relation with textural complexity and QPs. Generally speaking, if texture is very complex, CUs usually use large depths; and vice versa. In addition, if QPs are very large, CUs usually use small depths; and vice versa. Therefore, we can predict candidate depths based on textural complexity and QPs.

Suppose $v_{i,j}$ is a pixel value at (i, j) in a CU, n and \bar{v} are the size and average pixel value of a CU respectively, and the difference between the pixel value $v_{i,j}$ and the average value \bar{v} is denoted as $x_{i,j}$. Obviously, if the texture is very simple, the corresponding $x_{i,j}$ should be very small; and vice versa. Therefore, we can use $x_{i,j}$ to represent textural complexity. The expected value μ and the variance σ of $x_{i,j}$ are calculated by:

$$\mu = \frac{\sum_{i=0}^n \sum_{j=0}^n x_{i,j}}{n^2}, \sigma^2 = \frac{\sum_{i=0}^n \sum_{j=0}^n (x_{i,j} - \mu)^2}{n^2 - 1}. \quad (5)$$

The DCT coefficient at the position (i, j) is denoted as $x_F(i, j)$, its expected value and variance are denoted as $\mu_F(i, j)$ and $\sigma_F^2(i, j)$, respectively. They can be written as [19]:

$$\mu_F(i, j) = \mu \sqrt{[ARA^T]_{i,i} [ARA^T]_{j,j}} \\ \sigma_F^2(i, j) = \sigma^2 [ARA^T]_{i,i} [ARA^T]_{j,j}, \quad (6)$$

where $[ARA^T]_{i,i}$ is the component at (i, i) in a matrix, A is an 8×8 integer DCT transform matrix in SHVC, and R is:

$$R = \begin{bmatrix} 1 & \rho & \cdots & \rho^7 \\ \rho & 1 & \cdots & \rho^6 \\ \vdots & \vdots & \ddots & \vdots \\ \rho^7 & \rho^6 & \cdots & 1 \end{bmatrix}, \quad (7)$$

where ρ is the correlation coefficient and its value is set to 0.6 [24].

Since the distribution of $x_F(i, j)$ is uncertain, we can not obtain its probability based on its distribution. In order to address this issue, we can use Chebyshev's Inequality to estimate the probability of $x_F(i, j)$. According to the Chebyshev's Inequality, for any value ε , the following inequality holds:

$$p(|x_F(i, j) - \mu_F(i, j)| < \varepsilon) \geq 1 - \frac{\sigma_F^2(i, j)}{\varepsilon^2}. \quad (8)$$

When ε is greater than or equal to $3\sigma_F$, the probability in Eq. (8) is nearly higher than 90%, so we set ε as $3\sigma_F$. In this condition, the maximum absolute value of $x_F(i, j)$ is $3\sigma_F + |\mu_F|$. For Intra coding, in order to ensure any $x_F(i, j)$ can be quantized to zero, the following condition must be satisfied:

$$\frac{3\sigma_F + |\mu_F|}{Q_{step}} + \frac{1}{6} \leq 1. \quad (9)$$

where Q_{step} refers to quantization step. From Eq. (9), we have:

$$3\sigma_F + |\mu_F| \leq \frac{5}{6}Q_{step}. \quad (10)$$

Combining Eq. (6) and (10), we can derive:

$$\sigma \leq \left(\frac{5k_{i,j}}{6}Q_{step} - |\mu| \right) / 3, \quad (11)$$

where the value of $k_{i,j}$ is a constant value and is equal to $1/\sqrt{[ARA^T]_{i,i}[ARA^T]_{j,j}}$. The corresponding value of $k_{i,j}$ at the position (i, j) is listed below:

$$\begin{bmatrix} 0.3249 & 0.4248 & 0.5443 & 0.6966 & 0.8424 & 0.9673 & 1.0623 & 1.1206 \\ 0.4248 & 0.5554 & 0.7117 & 0.9108 & 1.1014 & 1.2647 & 1.3889 & 1.4652 \\ 0.5443 & 0.7117 & 0.9119 & 1.1671 & 1.4114 & 1.6206 & 1.7797 & 1.8775 \\ 0.6966 & 0.9108 & 1.1671 & 1.4937 & 1.8063 & 2.0741 & 2.2777 & 2.4028 \\ 0.8424 & 1.1014 & 1.4114 & 1.8063 & 2.1844 & 2.5083 & 2.7545 & 2.9058 \\ 0.9673 & 1.2647 & 1.6206 & 2.0741 & 2.5083 & 2.8802 & 3.1629 & 3.3367 \\ 1.0623 & 1.3889 & 1.7797 & 2.2777 & 2.7545 & 3.1629 & 3.4734 & 3.6642 \\ 1.1206 & 1.4652 & 1.8775 & 2.4028 & 2.9058 & 3.3367 & 3.6642 & 3.8655 \end{bmatrix}. \quad (12)$$

Combining Eq. (11) and (12), we can obtain the threshold values: if $\sigma \leq \left(\frac{5 \times 0.3249}{6}Q_{step} - |\mu| \right) / 3$, all quantized residual coefficients are zero, which is the condition of all-zero blocks; else if $\sigma \leq \left(\frac{5 \times 0.4248}{6}Q_{step} - |\mu| \right) / 3$, all quantized residual coefficients are zero except the coefficient at (0, 0); and so on. Accordingly, if $\sigma > \left(\frac{5 \times 3.8655}{6}Q_{step} - |\mu| \right) / 3$, all quantized residual coefficients may be not zero, which is the condition of all-nonzero blocks. Obviously, if all quantized residual coefficients are not zero, the current CU is not well predicted and can be early skipped. Conversely, if all quantized residual coefficients are zero, the current CU is very well predicted and can be early terminated. Therefore, we can use the condition of all-zero blocks and all-nonzero blocks to early terminate and early skip the current depth, respectively. Since depth levels in different CUs may have different probabilities, if we only use the textural based AZBANBs in prediction without considering depth probabilities, we definitely cannot obtain optimum performance. Therefore, we should combine depth probability with textural based AZBANBs to obtain threshold values for early skip and early termination.

3) *Combining Depth Probability with Textural Based AZBANBs in Prediction*: As both depth probability and the condition of textural based AZBANBs have obviously strong relationships with depth selection, we can combine them to predict candidate depth levels. As mentioned above, we can obtain each depth level's probability based on temporal and spatial correlations. We equally divide depth level probability into five categories: 0%-20%, 20%-40%, 40%-60%, 60%-80% and 80%-100%. These five categories can indicate different probabilities very well, e.g., 0%-20% indicates a very low possibility and 80%-100% indicates a very high possibility. Obviously, different categories should use different decision conditions for early skip and early termination. We can modify AZBANBs by multiplying the best values to obtain the best decision conditions for different categories. We define the best decision condition of a category as a coefficient value. Apparently, a coefficient value is not optimal for all probabilities in the corresponding category, but it can be considered as optimal for the median value of the corresponding category. Suppose l is the coefficient value of category 0% and 20%, it is not optimal for all probabilities in 0% and 20%, but it can be considered as optimal for 10% which is the median value of

category 0% and 20%. We integrate the coefficient values and their corresponding median values of all categories with LIP to derive the threshold values for depth early skip and depth early termination. Their corresponding processes are described as follows.

(a) Depth early skip

When depth level probability is in 0%-20%, our proposed method of selecting the coefficient value is as follow. We first use the condition of All-Nonzero Block mentioned above which is denoted as t_1 , then test with values greater than t_1 and smaller than t_1 . For values smaller than t_1 , we repeat dividing them by 2. While for values greater than t_1 , we repeat by adding $0.5t_1$. These values and their corresponding Bjøntegaard delta bit rates (BDBRs) [25] are listed in Table I. Given the same PSNR in the EL, BDBR refers to the bitrate difference compared with the SHVC reference software (SHM 11.0). A positive value and a negative value indicate coding efficiency decrease and increase in EL, respectively.

From Table I, we can observe that if a test value is larger than or equal to $0.5t_1$, the BDBRs in all sequences are 0.1% at most. This means that the coding efficiency loss is very small. Apparently, the smaller the value is, the more depths meet the condition. Therefore, we select $0.5t_1$ as the coefficient value (the best condition). We use the same way to obtain the coefficient values for the other depth probability categories in 20%-40%, 40%-60%, 60%-80% and 80%-100%, which are $1.0t_1$, $1.5t_1$, $2.0t_1$ and $3.5t_1$, respectively. Obviously, $0.5t_1$ is not the optimal value for all probabilities in 0%-20%, but it can be considered as the optimal value for the median value (10%) between 0 and 20%. In the same way, $1.0t_1$, $1.5t_1$, $2.0t_1$ and $3.5t_1$ also can be regarded as the optimal values for the median values of the other four categories, namely 30%, 50%, 70% and 90%, respectively. Putting these median values and their corresponding coefficient values together, we have five sets of values: (10%, $0.5t_1$), (30%, $1t_1$), (50%, $1.5t_1$), (70%, $2.0t_1$) and (90%, $3.5t_1$).

For any probability x , we use the five sets of values through LIP to calculate the corresponding predictive value dt :

$$dt = \sum_{i=0}^4 s_i \prod_{j=0, i \neq j}^4 \frac{x-x_j}{x_i-x_j}, \quad (13)$$

where x_0, x_1, x_2, x_3 and x_4 are 10%, 30%, 50%, 70% and 90%, respectively; s_0, s_1, s_2, s_3 and s_4 are $0.5t_1, 1.0t_1, 1.5t_1, 2.0t_1$ and $3.5t_1$, respectively. Based on Eqs. (13), we can derive the early skip condition below:

$$\sigma > \sum_{i=0}^4 s_i \prod_{j=0, i \neq j}^4 \frac{x-x_j}{x_i-x_j} \times \left(\frac{5 \times 3.8655}{6}Q_{step} - |\mu| \right) / 3. \quad (14)$$

(b) Early termination condition

In a similar way, we denote the condition of All-Zero Block mentioned above as t_2 and obtain the coefficient values for 0%-20%, 20%-40%, 40%-60%, 60%-80% and 80%-100%, which are $0.03125t_2, 0.0625t_2, 0.0625t_2, 0.125t_2$ and $0.125t_2$, respectively. Putting median values and their corresponding coefficient values of the five categories together, we also have five sets of values: (10%, $0.03125t_2$), (30%, $0.0625t_2$), (50%, $0.0625t_2$), (70%, $0.125t_2$) and (90%, $0.125t_2$). Based on the

TABLE I: Test values and corresponding BDBRs for depth probability in 0%-20%

Sequences \ Values	0.125 t_1	0.25 t_1	0.5 t_1	1 t_1	1.5 t_1	2 t_1
Blue_Sky	0.40%	0.30%	0.10%	0.00%	0.00%	0.00%
Ducks	0.20%	0.10%	0.00%	0.00%	0.00%	0.00%
Park_Joy	0.60%	0.30%	0.10%	0.00%	0.00%	0.00%
Pedestrian	0.30%	0.20%	0.00%	0.00%	0.00%	0.00%
Sunflower	0.10%	0.10%	0.00%	0.00%	0.00%	0.00%
Town	0.40%	0.20%	0.00%	-0.10%	-0.10%	-0.10%
Tractor	0.20%	0.10%	0.00%	0.00%	0.00%	0.00%
Tree	0.50%	0.20%	0.10%	0.00%	0.00%	0.00%

five sets of values, we use LIP to calculate the corresponding predictive value for any probability x using Eq. (13), and derive the early termination condition below:

$$\sigma \leq \sum_{i=0}^4 e_i \prod_{j=0, i \neq j}^4 \frac{x-x_j}{x_i-x_j} \times \left(\frac{5 \times 0.3249}{6} Q_{step} - |\mu| \right) / 3, \quad (15)$$

where x_0, x_1, x_2, x_3 and x_4 are 10%, 30%, 50%, 70% and 90%, respectively; e_0, e_1, e_2, e_3 and e_4 are 0.03125 t_2 , 0.0625 t_2 , 0.0625 t_2 , 0.125 t_2 and 0.125 t_2 , respectively.

Through the above process, we can obtain the early skip condition and early termination condition for depth selection. Since these conditions are developed only for 8×8 CUs, we can derive the corresponding conditions for other CU sizes. There are 64, 16 and 4 8×8 CUs in each 64×64, 32×32, 16×16 CU, respectively. We can multiply 64, 16 and 4 on the right sides of Eq. (14) and Eq. (15) to obtain the early skip condition and the early termination condition for the corresponding CU sizes.

B. Probability and Residual Coefficients-Based ILR Mode prediction (PRC-BIMP)

In Intra prediction of SHVC, each CU needs to check both Intra mode and ILR mode in EL. Since ILR mode is very simple and cost little time, in order to maintain coding efficiency, we always check ILR mode. As we know, if a mode is very well predicted, the residue follows a Gaussian distribution [15] or a Laplacian distribution [24]. We select Gaussian distribution in our experiments due to its superior performance [15]. Consequently, in order to improve coding speed, we can first check if the residue of ILR mode follows a Gaussian distribution, and then select the best condition based on experiments. In addition, a mode probability also strongly relates to mode selection. Based on the above analysis, we first obtain probabilities of ILR mode, and then combine the probabilities with residual coefficients to develop early termination condition for ILR mode.

1) *Correlation-Based Mode Prediction*: Similar to depth prediction, we combine neighboring CUs and their correlation degrees to obtain ILR mode probability. The neighboring CUs in EL are shown in Fig. 2. The mode of C, L, U, UL, UR, FC, FL, FU, FUL and FUR are denoted as $m_0, m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9$, respectively. Similar to depth prediction in the above section, the sum of absolute differences between neighboring CUs' modes in the current frame and their co-located CUs' modes in the previous frame is inversely related

to the temporal correlation degree of the current CU. Since the maximum value of the sum is 4, the temporal correlation degree of the current CU, tm , is calculated by:

$$tm = 4 - \sum_{i=1}^4 |m_i - m_{i+5}|, \quad (16)$$

where m_i is the mode of the i -th ($1 \leq i \leq 4$) neighboring CUs of the current CU in the current frame, and m_{i+5} is the mode of its co-located CU in the previous frame. We set ILR mode and Intra mode as 0 and 1, respectively.

Similar to depth prediction, we also use the spatial correlation degrees of CUs in the previous frame as those of the co-located CUs in the current frame. For example, we use the spatial correlation degrees of CU FL and CU FC as those of CU L and CU C. Suppose \mathbf{sm} is the spatial correlation degree vector of four neighboring CUs' modes of the CU FC in the previous frame. It is obvious that the absolute difference between the mode of the CU FC and its i -th neighboring CU in the previous frame, $|m_{i+5} - m_5|$, is inversely related to their corresponding spatial correlation degree. Since the maximum value of a mode is 1, we can calculate the i -th component of the spatial correlation degree vector, sm_i , by:

$$sm_i = 1 - |m_{i+5} - m_5|, \quad (17)$$

where m_5 is the mode of CU FC and m_{i+5} is the mode of its i -th neighboring CU in the previous frame. Through the above process, we can obtain the temporal and spatial correlation degrees of the current CU's mode. Assume \mathbf{nm} is the mode vector of four neighboring CUs of the current CU, and its i -th component nm_i 's value is m_i . Combining modes of relative CUs with their correlation degrees, we use the below Bayesian formulation to derive the probability of mode m_0 , $f_m(m_0)$, as:

$$\begin{aligned} f_m(m_0) &= p(m_0 | ((nm, sm), (m_5, tm))) = \frac{p(((nm, sm), (m_5, tm)), m_0)}{p((nm, sm), (m_5, tm))} \\ &= \frac{p(((nm, sm), (m_5, tm)) | m_0) p(m_0)}{p((nm, sm), (m_5, tm))}. \end{aligned} \quad (18)$$

Similar to depth prediction, we also use a Naive Bayesian classifier to calculate the probability of mode m_0 by:

$$f_m(m_0) = \frac{p((m_5, tm) | m_0) p(m_0) \prod_{i=0}^3 p((nm_i, sm_i) | m_0)}{p(m_5, tm) \prod_{i=0}^3 p(nm_i, sm_i)}. \quad (19)$$

In order to calculate $f_m(m_0)$, we need to obtain $p(m_0)$, $p(\mathbf{m}_5, \mathbf{tm})$, $p((\mathbf{m}_5, \mathbf{tm}) | \mathbf{m}_0)$, $p(\mathbf{nm}_i, \mathbf{sm}_i)$ and $p((\mathbf{nm}_i, \mathbf{sm}_i) | \mathbf{m}_0)$ first through extensive experiments, and then derive the probabilities of ILR and Intra modes.

2) *Combining Mode Probability with Residual Coefficients in Prediction*: As mentioned above, if a mode is very well predicted, the residue should follow a Gaussian distribution. Thus, we use Jarque-Bera test to determine whether its residue follows a Gaussian distribution.

Suppose r_0, r_1, \dots, r_n are residual coefficients, their skewness S and kurtosis K are:

$$S = \frac{B_3}{B_2^{3/2}}, K = \frac{B_4}{B_2^2}, \quad (20)$$

where B_k ($k = 2, 3, 4$) is the sample central moment of order k and is derived by:

$$B_k = \frac{\sum_{i=1}^n (r_i - \bar{r})^k}{n}, \quad (21)$$

where \bar{r} and n are the average value and the number of residual coefficients, respectively. Jarque-Bera Statistics, JB , is:

$$JB = \frac{S^2}{6/n} + \frac{(K-3)^2}{24/n}. \quad (22)$$

According to the knowledge of mathematical statistics, JB follows a chi-squared distribution with the degree of freedom of 2.

Since both a mode probability and its residual coefficients have strong relationships with mode selection, we combine ILR mode probability with its residual coefficients to decide whether the ILR mode is the best mode. Similar to depth prediction, we also evenly divide ILR mode probability into five groups. For the ILR mode probability in 0%-20%, we use some common significance level values in testing. According to chi-squared distribution table, for significance level values at 0.95, 0.90, 0.80, 0.70, 0.50, 0.30, 0.20, 0.10, 0.05, 0.025 and 0.01, the corresponding test critical values are 0.10, 0.21, 0.45, 0.71, 1.39, 2.41, 3.22, 4.60, 5.99, 7.38 and 9.21, respectively. These test critical values and the corresponding BDBRs are listed in Table II.

From Table II, if a test critical value is equal to or less than 0.21, the BDBRs in all sequences are 0.1% at most. This means that the coding efficiency loss is very small. Of course, if we select larger test values, the corresponding BDBRs will also be larger. Considering the tradeoff between coding speed and coding efficiency, we select 0.21 as the coefficient value (the best value). In the same way, we can obtain the coefficient values for 20%-40%, 40%-60%, 60%-80% and 80%-100%, which are 1.39, 2.41, 3.22 and 4.6, respectively. Putting the median values and their corresponding coefficient values of the above five groups together, we have five sets of values: (10%, 0.21), (30%, 1.39), (50%, 2.41), (70%, 3.22) and (90%, 4.6). For any ILR mode probability x , we use the five sets of values through LIP to derive the corresponding predictive value mt :

$$mt = \sum_{i=0}^4 r_i \prod_{j=0, i \neq j}^4 \frac{x - x_j}{x_i - x_j}, \quad (23)$$

where x_0, x_1, x_2, x_3 and x_4 are 10%, 30%, 50%, 70% and 90%, respectively; r_0, r_1, r_2, r_3 and r_4 are 0.21, 1.39, 2.41, 3.22 and 4.6, respectively. Combining with Eqs. (22) and (23), we can derive the decision condition for ILR mode as:

$$JB \leq \sum_{i=0}^4 r_i \prod_{j=0, i \neq j}^4 \frac{x - x_j}{x_i - x_j}. \quad (24)$$

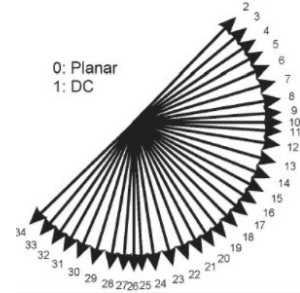


Fig. 3: DMs in SHVC.

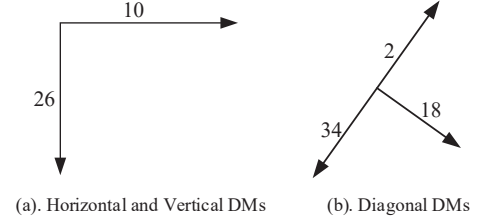


Fig. 4: Two sets of Typical DMs.

If Eq. (24) is satisfied, the ILR mode can be considered as the best mode, and then we can directly skip Intra mode prediction.

C. Probability and HC-Based DM Prediction (PH-BDP)

Similar to HEVC, Intra prediction in SHVC includes 33 DMs, Direct Current (DC) and Planar modes for Luma components, as shown in Fig. 3. SHVC firstly uses the RMD process to obtain the first N modes with the smallest HCs, and then uses the RDO process to select the mode with the smallest RD cost. For large CUs, e.g., 64×64 , 32×32 , 16×16 , N is 3; while for small CUs, like 8×8 and 4×4 , N is 8. In order to improve coding speed, we develop several methods below.

1) The Difference of HC-Based DM Prediction in RMD:

In order to obtain candidate DMs, we fully use the differences of HCs of typical DMs in prediction. In Fig. 3, DM 10 and 26 are Horizontal and Vertical DMs. DM 2, 18 and 34 are Diagonal DMs. Fig. 4 shows two sets of typical DMs, including Horizontal and Vertical DMs (Fig. 4 (a)) and Diagonal DMs (Fig. 4 (b)).

It is obvious that the HCs of DMs are strongly related with DM selection. Therefore, we propose to first check the above two typical DM sets, and then predict candidate DMs based on the differences of their HCs. The details of the proposed candidate DM prediction are described below. The process is summarized in Fig. 5, in which HDM denotes HC of DM and “ \ll ” represents “significantly smaller than”.

- (1) Since DM 0 and 1 are often selected as the best DMs, we first test these two DMs, and then test DMs 10, 26, 2, 18 and 34. If the HCs of all these DMs are the same, go to (2). Otherwise, if the HC of DM 10 is significantly smaller than that of DM 26, go to (3); else if the HC of DM 26 is significantly smaller than that of DM 10, go to (4); else if the HC of DM 18 is significantly smaller than those of DMs 2 and 34, go to (5); else if the HC

TABLE II: Test critical values and corresponding BDBRs for ILR mode probability in 0%-20%

Sequences \ Values	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.6	5.99	7.38	9.21
Blue_Sky	0.00	0.00	0.00	0.00	-0.10	-0.10	-0.10	-0.10	-0.20	-0.20	-0.20
Ducks	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Park_Joy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.10
Pedestrian	0.00	0.00	0.00	0.00	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10
Sunflower	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.10	-0.10	-0.10	-0.10
Town	0.00	0.10	0.20	0.30	0.50	0.90	1.00	1.30	1.50	1.60	1.70
Tractor	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.10	-0.10	-0.10	-0.10
Tree	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

of DM 18 is significantly larger than those of DMs 2 and 34, go to (6); else, go to (7).

- (2) The reference pixels of the current CU should be the same. Therefore, the HCs of all the 35 DMs should be the same and we only select DM 0 to check.
- (3) DMs 2-18 are very likely to be selected. Since DM 2 represents horizontal upper right and DM 18 represents lower right, we further predict candidate DMs based on their HCs: if the HC of DM 2 is significantly smaller than that of DM 18, DMs 3-9 will be checked further; if the HC of DM 18 is significantly smaller than that of DM 2, DMs 11-17 will be checked further; else DMs 3-9 and DMs 11-17 will be checked further.
- (4) DMs 18-34 are very likely to be selected. We use a similar way like (3) to check the corresponding candidate DMs.
- (5) DMs 11-17 and DMs 19-25 are selected to check further.
- (6) DM 2 and DM 34 have the same angle but in the opposite directions, so they have different reference pixels. If the HC of DM 2 is significantly smaller than that of DMs 34, DMs 3-9 are selected to check; else if the HC of DM 34 is significantly smaller than that of DMs 2, DMs 27-33 are selected to check; else DMs 3-9 and DMs 27-33 are selected to further check.
- (7) All other DMs are selected to check further.

Through the above process, we can obtain candidate DMs. Here, the key is how to obtain the significantly smaller condition.

Suppose the CU residual matrix of two DMs are R_1 and R_2 , their difference R is:

$$R = R_1 - R_2. \quad (25)$$

Its Hadamard transform value is:

$$HRH = HR_1H - HR_2H, \quad (26)$$

where H is a 4×4 Hadamard matrix. According to Cauchy-Inequality, we can derive:

$$\begin{aligned} HRH &\leq \left| \sum_{i=0}^4 \sum_{j=0}^4 (HH^T)^2 \right|^{\frac{1}{2}} \times \left| \sum_{i=0}^4 \sum_{j=0}^4 r^2(i, j) \right|^{\frac{1}{2}} \\ &\leq 2 \left| \sum_{i=0}^4 \sum_{j=0}^4 r_{ij}^2 \right|^{\frac{1}{2}} \leq 2 \sum_{i=0}^4 \sum_{j=0}^4 |r_{ij}|. \end{aligned} \quad (27)$$

Suppose x_{ij} is the value at the location (i, j) in HRH , we can derive:

$$x_{ij} = \sum_{k=0}^4 \sum_{p=0}^4 h_{ik} r_{kp} h_{pj} \leq \sum_{k=0}^4 \sum_{p=0}^4 |h_{ik} h_{pj}| |r_{kp}| \leq \sum_{k=0}^4 \sum_{p=0}^4 |r_{kp}|. \quad (28)$$

For R_1 and R_2 , if any quantized values in HRH are not equal to 0, they are significantly different. In order to ensure that any quantized values in HRH are not equal to 0, according to Eq. (28), we can derive:

$$\sum_{k=0}^4 \sum_{p=0}^4 |r_{kp}| > Q_{step}. \quad (29)$$

Combining Eq. (26), (27), (28), (29), we can derive significantly smaller condition:

$$|HR_1H - HR_2H| > 2Q_{step}. \quad (30)$$

Eq. (30) can decide if HR_2H is significantly smaller than HR_1H . In order to decide if one is significantly smaller than the other one, we need to rewrite the above condition. In addition, Eq. (30) is developed only for 4×4 CUs. For other CU sizes, their corresponding significantly smaller condition can be rewritten below:

$$HC_1 - HC_2 > 2mQ_{step}, \quad (31)$$

where m is the number of 4×4 CUs in the current CU, which is 256, 64, 16, and 4 for CU size 64×64 , 32×32 , 16×16 , and 8×8 , respectively. If Eq. (31) is satisfied, we can consider that HC_2 is significant smaller than HC_1 .

2) *Gradient Based Early Termination in RMD*: Gradient has been widely used for DM prediction [26]. We can use a Sobel operator to obtain the direction and amplitude of each DM. Undoubtedly, DMs with larger amplitudes are more likely to be selected as the best DM, and vice versa. We use a classical Sobel operator to calculate the gradient of the pixel at (i, j) as follows:

$$G_{i,j}^x = p_{i-1,j+1} + 2 \times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i,j-1} - p_{i+1,j-1}, \quad (32)$$

$$G_{i,j}^y = p_{i+1,j-1} + 2 \times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i-1,j} - p_{i-1,j+1}, \quad (33)$$

where $G_{i,j}^x$ and $G_{i,j}^y$ refer to the differences of vertical and horizontal directions respectively. The amplitude of the gradient is:

$$Amp(G_{i,j}) = |G_{i,j}^x| + |G_{i,j}^y|. \quad (34)$$

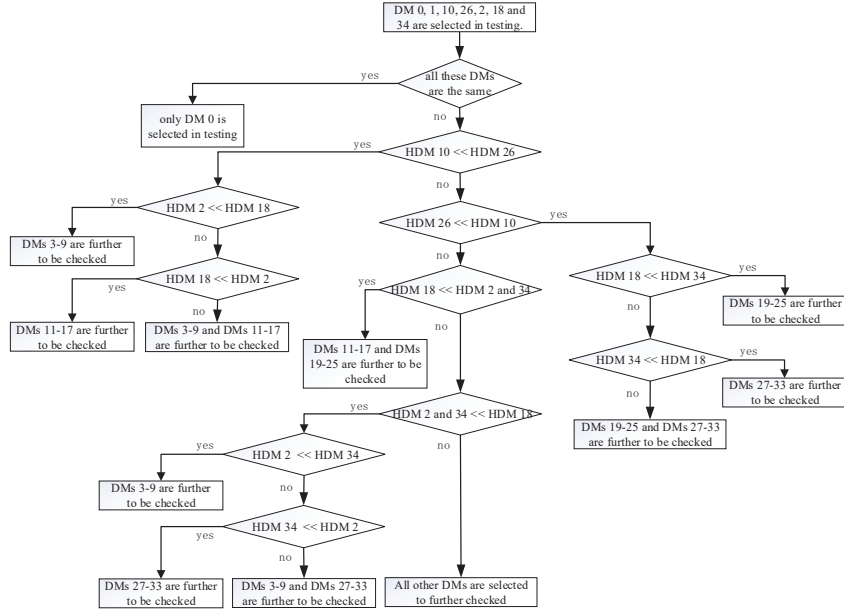


Fig. 5: Flowchart of the DM Prediction algorithm.

The direction of the gradient is calculated by:

$$\text{Ang}(G_{i,j}) = \arctan\left(\frac{G_{i,j}^y}{G_{i,j}^x}\right). \quad (35)$$

In order to reduce computational complexity, we use $G_{i,j}^y / G_{i,j}^x$ instead of \arctan . Using the above process, the amplitude and the direction of each gradient can be obtained. Subsequently, the amplitudes of all gradients closest to the direction of each DM are separately summed up. Through the above process, we can obtain the amplitudes of all DMs and sort DMs by their amplitudes in descending order. However, the amplitudes of DMs cannot totally reflect the probabilities of corresponding DMs. In order to address the issue, we use the percentage of amplitudes of DMs instead. Suppose a_i is the amplitude of the i -th DM, the percentage of the first k DMs is:

$$p_k = \frac{\sum_{i=0}^k a_i}{\sum_{i=0}^{33} a_i}. \quad (36)$$

In addition, DMs with smaller HCs are also more likely to be selected as the best DM, and vice versa. Suppose R is the residual matrix of the current CU, if any quantized value in HRH is 0, the corresponding DM is very likely to be the best DM. Using the same way as above, we can derive the early termination condition for DM selection below:

$$HC < 2mQ_{step}, \quad (37)$$

where m is the number of 4×4 CUs in the current CU.

Combining the amplitude percentage and HCs, we modify Eq. (37) to derive the corresponding condition below:

$$HC < t \times 2mQ_{step}, \quad (38)$$

where t is a parameter to adjust the above condition. For a DM with a larger percentage, we can set a larger t to improve

coding speed. Conversely, we can set a smaller t to maintain coding efficiency. Similar to the depth prediction process, we evenly divide the percentage into five groups, namely, **0%-20%**, 20%-40%, 40%-60%, 60%-80% and 80%-100%, and the corresponding coefficient values (the best values) are $2m$, $4m$, $8m$, $10m$ and $14m$, respectively. Similarly, for the median values of the five groups above, 10%, 30%, 50%, 70% and 90%, their corresponding coefficient values are $2m$, $4m$, $8m$, $10m$ and $14m$, respectively. Putting these median values and their corresponding coefficient values together, we obtain five sets of values: (10%, $2m$), (30%, $4m$), (50%, $8m$), (70%, $10m$) and (90%, $14m$). We use these five sets of values through LIP to calculate the corresponding predictive value for any percentage x , and combine with Eq. (38) to derive the early termination below:

$$HC < 2mQ_{step} \times \sum_{i=0}^4 l_i \prod_{j=0, i \neq j}^4 \frac{x-x_j}{x_i-x_j}, \quad (39)$$

where x_0, x_1, x_2, x_3 and x_4 are 10%, 30%, 50%, 70% and 90%, respectively; l_0, l_1, l_2, l_3 and l_4 are 1, 2, 4, 5 and 7, correspondingly.

3) *HC-Based Early Termination in RDO*: Through the above process, we can obtain the first N DMs with the smallest HCs, and then select the DM with the smallest RD cost in RDO process. If the HCs of two neighboring DMs are very similar, their corresponding RD costs should also be very similar, and thus we only need to check either one of them. Otherwise, if the HCs of two neighboring DMs are significantly different, their corresponding RD costs should also be different, and it is unlikely to adopt the DM with a larger HC. Therefore, if two neighboring DMs are very similar or significantly different, we only need to check the DM with a smaller HC. Here, we revise the significantly different condition (30) and similar condition (37) through experiments,

TABLE III: Performances of different strategies

Sequence	PTC-BDP		PRC-BIMP		PH-BDP	
	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	0.10%	42.96%	0.0%	45.82%	-0.1%	32.22%
PeopleOnStreet	0.10%	40.48%	-0.2%	42.65%	-0.1%	33.79%
Kimono	-0.20%	45.24%	-0.2%	59.21%	-0.1%	31.96%
ParkScene	0.00%	42.82%	-0.1%	41.81%	-0.1%	32.09%
Cactus	0.30%	42.22%	0.5%	48.09%	0.3%	32.33%
BasketballDrive	0.70%	46.77%	1.7%	50.45%	0.5%	32.43%
BQTerrace	0.40%	40.93%	1.1%	46.62%	0.2%	34.09%
Average	0.20%	43.06%	0.40%	47.81%	0.09%	32.70%

and derive the similar condition below:

$$|HC_1 - HC_2| < mQ_{step}. \quad (40)$$

The significant different condition is:

$$|HC_1 - HC_2| > 5mQ_{step}. \quad (41)$$

If conditions in Eqs. (40) or (41) are satisfied, we only need to check the one with a smaller HC. Otherwise, we need to check both of them. In addition, if the HCs of two neighboring DMs are very similar, it is very likely that these two DMs include the best DM, so we do not need to further check other DMs. If the HCs of two neighboring DMs are significantly different, the DM with a larger HC is unlikely to be adopted and subsequent DMs are even more unlikely to be adopted, so we also do not need to further check other DMs.

V. EXPERIMENTAL RESULTS

In order to verify the performance of the proposed fast Intra prediction algorithm for spatial SHVC, the proposed algorithm has been implemented on the reference software (SHM 11.0) and tested on a server with Intel (R) 2.0 GHz CPU and 30 GB memory. In order to demonstrate the generality of our proposed algorithm, training and testing sequences do not overlap. Coding efficiency and coding speed are used to evaluate the performance of the proposed algorithm. Coding efficiency is indicated by BDBR [25]. Coding speed is denoted by TS, which evaluates the percentage of encoding run-time savings in EL only.

The proposed algorithm includes three strategies: “PTC-BDP”, “PRC-BIMP” and “PH-BDP”. As mentioned above, we also only test the performance under the scalability ratio 2x and the QP set (24, 28, 32, 36) in EL. The performances of the three strategies are shown in Table III.

Since each depth prediction process includes both ILR prediction and Intra prediction, “PTC-BDP” can improve the coding speed very significantly. The proportion of ILR mode is significantly larger than that of Intra mode, and its coding process is very simple. “PRC-BIMP” can determine if ILR mode is the best mode so as to skip Intra mode prediction in many CUs, and can improve the coding speed very remarkably. Based on the proportions of gradient amplitudes and HCs of DMs, “PHC-BDMP” can predict candidate DMs and can early terminate both RMD and RDO processes, and thus can also speed up coding effectively.

In order to further demonstrate the performance of the proposed algorithm, we compare its performance, which integrates all of the three proposed strategies, with that of EETBS

TABLE IV: Performance comparison of different methods (case 1)

Sequence	Proposed		EETBS [12]		FIICA [14]	
	BDBR	TS	BDBR	TS	BDBR	TS
Kimono	-0.42%	68.16%	0.39%	71.43%	-0.21%	62.35%
ParkScene	-0.22%	61.13%	0.01%	62.86%	-0.12%	38.17%
Cactus	-0.12%	59.73%	-0.22%	46.05%	-0.18%	41.89%
BasketballDrive	0.06%	63.74%	0.48%	48.48%	0.40%	47.06%
BQTerrace	-0.02%	56.91%	0.30%	47.89%	0.41%	46.27%
Average	-0.14%	61.93%	0.20%	55.34%	0.06%	47.15%

TABLE V: Performance comparison of different methods (case 2)

Sequence	Proposed		EETBS [12]		FIICA [14]	
	BDBR	TS	BDBR	TS	BDBR	TS
Kimono	-0.68%	69.72%	0.60%	72.69%	0.81%	61.67%
ParkScene	-0.68%	64.17%	-1.10%	65.13%	-1.22%	37.46%
Cactus	-0.38%	62.53%	0.42%	45.83%	0.31%	40.13%
BasketballDrive	-0.23%	66.17%	-1.01%	49.32%	-0.80%	44.36%
BQTerrace	-0.12%	59.82%	0.09%	48.53%	0.00%	45.15%
Average	-0.42%	64.48%	-0.20%	56.30%	-0.18%	45.75%

TABLE VI: Performance comparison of different methods (case 3)

Sequence	Proposed		EETBS [12]		FIICA [14]	
	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	-0.01%	56.00%	0.30%	53.12%	0.41%	36.37%
PeopleOnStreet	-0.08%	54.60%	0.01%	51.91%	0.10%	39.43%
Kimono	-0.19%	66.13%	-0.13%	70.12%	-0.13%	60.27%
ParkScene	-0.10%	57.40%	0.21%	63.87%	0.22%	36.49%
Cactus	0.68%	56.20%	0.81%	45.63%	0.89%	37.92%
BasketballDrive	1.12%	60.60%	0.80%	47.59%	0.71%	41.48%
BQTerrace	0.49%	53.62%	0.40%	49.16%	0.50%	43.56%
Average	0.27%	57.79%	0.34%	54.49%	0.38%	42.22%

TABLE VII: Performance comparison of different methods (case 4)

Sequence	Proposed		EETBS [12]		FIICA [14]	
	BDBR	TS	BDBR	TS	BDBR	TS
Traffic	-0.15%	60.28%	-0.41%	53.47%	-0.30%	37.89%
PeopleOnStreet	-0.20%	56.74%	-0.33%	52.38%	-0.23%	40.15%
Kimono	-0.39%	67.38%	0.31%	70.83%	0.19%	60.18%
ParkScene	-0.10%	60.77%	0.10%	64.79%	0.11%	38.13%
Cactus	0.54%	59.19%	0.51%	43.17%	0.70%	39.29%
BasketballDrive	1.08%	63.34%	1.50%	46.73%	1.72%	42.74%
BQTerrace	0.47%	54.90%	0.42%	47.43%	0.61%	44.37%
Average	0.18%	60.37%	0.30%	54.11%	0.40%	43.25%

algorithm [12] and FIICA algorithm [14]. To the best of our knowledge, these two algorithms are the best and most recent algorithms for spatial Intra SHVC. For fair comparisons, all algorithms are tested on the same computing platform. As mentioned above, due to two scalability ratios and two QPs settings, we classify their combinations into four cases in EL. Case 1 is scalability ratio 1.5x and the QP set (22, 26, 30, 34) in EL, case 2 is scalability ratio 1.5x and the QP set (24, 28, 32, 36) in EL, case 3 is scalability ratio 2x and the QP set (22, 26, 30, 34) in EL, and case 4 is scalability ratio 2x and QP set (24, 28, 32, 36) in EL. The overall performance comparisons in terms of coding efficiency and coding speed are listed in Table IV (case 1), Table V (case 2), Table VI (case 3) and Table VII (case 4), respectively. According to CSTC, sequences with 1.5x only need to test 5 sequences while sequences with 2x need to test 7 sequences. Therefore, there are only five sequences for Tables IV and V, and seven sequences for TABLE VI and TABLE VII.

From Table IV to Table VII, for all the above four cases, we

TABLE VIII: Overall average performance comparison of the different methods

Case	Proposed		EETBS [7]		FIICA [9]	
	BDBR	TS	BDBR	TS	BDBR	TS
Case 1	-0.14%	61.93%	0.20%	55.34%	0.06%	47.15%
Case 2	-0.42%	64.48%	-0.20%	56.30%	-0.18%	45.75%
Case 3	0.27%	57.79%	0.34%	54.49%	0.38%	42.22%
Case 4	0.18%	60.37%	0.30%	54.11%	0.40%	43.25%
Average	-0.03%	61.14%	0.16%	55.06%	0.17%	44.60%

can find that the proposed algorithm significantly outperforms those two reference algorithms, in terms of coding efficiency (BDBR) as well as coding speed (TS).

In order to clearly demonstrate the performance of the proposed algorithm, Table VIII provides the overall average performance comparisons among these three algorithms with all four cases. In Table VIII, the overall average BDBRs of the proposed algorithm, EETBS, and FIICA are -0.03%, 0.16% and 0.17%, respectively. While the overall average TS of the proposed algorithm, EETBS, and FIICA are 61.14%, 55.06% and 44.60% correspondingly. Compared with the other two algorithms, we can observe that the average BDBR of the proposed algorithm is smaller, meanwhile the average TS of the proposed algorithm is significantly faster. Therefore, we can conclude that the proposed algorithm performs much better than both reference algorithms, not only on coding speed but also on coding efficiency.

The main reasons why the proposed algorithm can effectively improve the coding speed are: (1) more features, including correlations, correlation degrees, textural complexity and QPs, are used together through LIP to obtain more accurate depth decision condition, thus more depths can be early skipped or early terminated; (2) more attributes, including correlations, correlation degrees and residual coefficient distribution, are jointly exploited through LIP to derive more accurate decision condition for ILR mode, so more CUs can skip Intra encoding process; (3) more features, including the differences of HCs of typical DMs, the percentages of gradient amplitudes, HCs and the differences of HCs of neighboring IMs, are adopted, thus more DMs are excluded in both RMD process and RDO process. Through the above process, many depths, ILR modes and DMs can be excluded in prediction, and then coding speed can be improved significantly.

Generally speaking, the improvement of coding speed will lead to an increase in BDBR, namely a decrease in coding efficiency. However, from Table IV to Table VIII, we can sometimes observe that coding efficiency increases, i.e., BDBR savings, on some sequences, when compared against the SHM reference software. One major reason is due to the Intra prediction process, in which CUs are predicted by their reference pixels, shown in black dots in Fig. 6. Apparently, if the texture of CUs and their reference pixels are more similar, their corresponding Intra prediction will be more accurate and RD costs will be smaller accordingly, and vice versa. For the current CU, different methods will lead to different neighboring CUs, which will lead to different reference pixels for the current CU, and in turn, lead to different RD costs for the current CU. In other words, compared to SHM, our

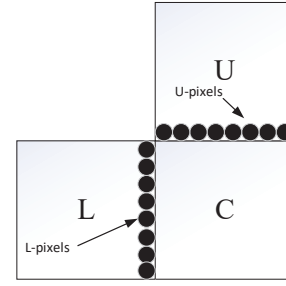


Fig. 6: Intra prediction in CTU

method may therefore occasionally achieve either a decrease or an increase in BDBR [18]. Combined with some effective and efficient bit allocation schemes, coding efficiency may be improved [27-28].

VI. CONCLUSION

In this paper, we have proposed a new probability-based Intra prediction algorithm for spatial SHVC. One unique feature is that different kinds of probabilities are fully exploited in prediction. In order to improve coding speed, we have developed three fast strategies corresponding to depth prediction, ILR mode early termination and DM prediction. Different depths, ILR modes and DMs may have different probabilities, which have strong relationships with their individual predictions. Therefore, in this paper, we integrate multiple types of probabilities to obtain accurate prediction and decision conditions. Experiments demonstrate that the proposed algorithm can achieve superior performances than other relative algorithms. The proposed algorithm allows for cost effective simulcast transmission of the same video with different screen resolutions in one bit-stream. Therefore, it is very useful for broadcasters.

REFERENCES

- [1] 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.
- [2] J. Lei, D. Li, Z. Pan, Z. Sun, S. Kwong, and C. Hou, "Fast Intra Prediction Based on Content Property Analysis for Low Complexity HEVC-Based Screen Content Coding," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 48–58, Mar. 2017.
- [3] Y. T. Kuo, P. Y. Chen, and H. C. Lin, "A Spatiotemporal Content-Based CU Size Decision Algorithm for HEVC," *IEEE Trans. Broadcast.*, vol. 66, no. 1, pp. 100–112, Mar. 2020.
- [4] H. L. Tan, C. C. Ko, and S. Rahardja, "Fast coding quad-tree decisions using prediction residuals statistics for high efficiency video coding (HEVC)," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 128–133, Mar. 2016.
- [5] Y. Li, G. Yang, Y. Song, H. Zhang, D. Zhang, "Early Intra CU Size Decision for Versatile Video Coding Based on a Tunable Decision Model," *IEEE Trans. Broadcast.* (10.1109/TBC.2021.3073556)
- [6] M. Jamali, and S. Coulombe, "Fast HEVC Intra Mode Decision Based on RDO Cost Prediction," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 109–122, Mar. 2019.
- [7] 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.*, Vancouver, BC, Canada, May 2013, pp. 1744–1748.

- [8] 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. Informat. BCI, Thessaloniki, Greece, 2013, pp. 61–65.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] X. Lu, C. Yu and G.R.Martin, "Efficient early termination and bypass scheme for CU size decision in spatially scalable SHVC intra coding," *Electronics Letters*, vol. 54, no. 22, pp. 1274–1276, Nov. 2018.
- [13] X. Lu, C. Yu, Y.F. Gu, and G. Martin, "A fast Intra coding algorithm for spatial scalability in SHVC," in Proc. IEEE Int. Conf. Image Processing (ICIP), Athens, Greece, Oct. 2018, pp. 1792–1796.
- [14] X. Lu, C. Yu and G.R.Martin, "Fast intra-and inter-coding algorithms for the spatially scalable extension of H.265/HEVC," *Multimedia Tools and Applications*(10.1007/s11042-020-09085-0)
- [15] D. Wang, C. Zhu, Y. Sun, F. Dufaux, and 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.
- [16] D. Wang, Y. Sun, C. Zhu, W. Li, and F. Dufaux, "Fast depth and inter mode prediction for quality scalable high efficiency video coding," *IEEE Trans. Multimedia.*, vol. 22, no. 4, pp. 833–845, Apr. 2020.
- [17] D. Wang, Y. Sun, W. Li, C. Zhu, and F. Dufaux, "Fast inter mode predictions for SHVC," in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Shanghai, China, Jul. 2019, pp. 1696–1701.
- [18] D. Wang, Y. Sun, C. Zhu, W. Li, F. Dufaux, and J. Luo, "Fast Depth and Mode Decision in Intra Prediction for Quality SHVC," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 6136 - 6150, Dec. 2020.
- [19] L. Shen and G. Feng, "Content-based adaptive SHVC mode decision algorithm," *IEEE Trans. Multimedia.*, vol.21, no. 11, pp. 2714 -2725, Apr. 2019.
- [20] L. Shen, G. Feng and P. An, "SHVC CU Processing Aided by a Feedforward Neural Network", *IEEE Trans. Industrial Informatics.*, vol. 15, no. 11, pp. 5803-5815, Apr. 2019.
- [21] X. G. Liu, Y. Liu, P.C Wang, C.F Lai, and H.C Chao, "An Adaptive Mode Decision Algorithm Based on Video Texture Characteristics for HEVC Intra Prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 8, pp. 1737-1748, Aug. 2017.
- [22] 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.
- [23] Common SHM Test Conditions and Software Reference Configurations, document JCTVC-Q1009, ITU-T SG 16 WP 3 ISO/IEC JTC1/SC29/WG., Mar. 2014.
- [24] I.-M. Pao and M.-T. Sun, "Modeling DCT coefficients for fast video encoding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 608–616, Jun. 1999.
- [25] G. Bjontegaard, "Calculation of average PSNR difference between RD-curves," in Proc. 13th VCEG-M33 Meeting, Austin, TX, USA, Apr. 2001, pp. 2-4.
- [26] M. Jamali, S. Coulombe, and F. Caron, "Fast HEVC intra mode decision based on edge detection and SATD costs classification," in *Proc. Data Compress. Conf.*, Snowbird, UT, USA, Apr. 2015, pp. 43-52.
- [27] H. Guo, C. Zhu, S. Li, Y. Gao, "Optimal Bit Allocation at Frame Level for Rate Control in HEVC," *IEEE Trans. Broadcast.*, vol. 65, no. 2, pp. 270-281, Jul. 2018.
- [28] H. Guo, C. Zhu, M. Xu, S. Li, "Inter-Block Dependency-Based CTU Level Rate Control for HEVC," *IEEE Trans. Broadcast.*, vol. 66, no. 1, pp. 113 - 126, Mar. 2020.



University of Electronic Science and Technology of China. His main research interest is video coding.

Dayong Wang received the Ph.D. degree in computer science from the University of Electronic Science and Technology of China in 2010. He was a Lecturer with the Hubei University of Arts and Science from 2010 to 2012. He held a postdoctoral research position with the Graduate School, Shenzhen, Tsinghua University, from 2012 to 2015. He is currently an Associate Professor with the Chongqing University of Posts and Telecommunications, and an adjunct professor with Shangrao Normal University, and also a part-time Post-Doctoral Fellow with the



refereed journals and conferences. She is a member of the IEEE Multimedia Communications Technical Committee and was a Co-Editor of the IEEE Distributed Systems Online. She has served as a session chair, a member of organizing committees and technical program committees for over 55 international conferences.

Yu Sun received the B.S. and M.S. degrees in computer science from the University of Electronic Science and Technology of China, and the Ph.D. degree in computer science and engineering from the University of Texas at Arlington in 2004. She is currently a Professor with the Department of Computer Science, the University of Central Arkansas, USA. Her main research interests include video compression, multimedia communication, and image processing. Her research work in these areas has been published in over 70 technical papers in



at UESTC. Since 2020, he has been a professor at the School of Mathematical and Computer Sciences, Shangrao Normal University. His research interests include image processing, video analysis, and deep learning.

Jinhua Liu Jinhua Liu received the B.S. and M.S. degrees from the Jiangxi Normal University and the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2003 and 2008, respectively, both in School of Computer Science and Engineering, and the Ph.D degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2011. From 2012 to 2016, he was a senior engineer at Sichuan JiuZhou Electric Group Co., Ltd. Since 2015, He is pursuing the postdoctoral in School of Electronic Engineering



Frederic Dufaux (S'93-M'95-SM'09-F'16) is a CNRS Research Director at Laboratoire des Signaux et Systèmes (L2S, UMR 8506), CNRS-CentraleSupélec-Université Paris-Sud, where he is head of the Telecom and Networking division.

Frederic received his M.Sc. in physics and Ph.D. in electrical engineering from EPFL in 1990 and 1994 respectively. He has over 20 years of experience in research, previously holding positions at EPFL, Emitall Surveillance, Genimedia, Compaq, Digital Equipment, and MIT. He was also Editor-

in-Chief of Signal Processing: Image Communication from 2010 until 2019.

Frederic is a Fellow of IEEE. He was Chair of the IEEE SPS Multimedia Signal Processing (MMSP) Technical Committee in 2018 and 2019. He is a member of the IEEE SPS Technical Directions Board. He was Vice General Chair of ICIP 2014, General Chair of MMSP 2018, and Technical Program co-Chair of ICIP 2019. He will be Technical Program co-Chair of ICIP 2021. He is also a founding member and the Chair of the EURASIP Technical Area Committee on Visual Information Processing. Frederic has been involved in the standardization of digital video and imaging technologies for more than 15 years, participating both in the MPEG and JPEG committees. He was co-chairman of JPEG 2000 over wireless (JPWL) and co-chairman of JPSearch. He is the recipient of two ISO awards for these contributions. His research interests include image and video coding, 3D video, high dynamic range imaging, visual quality assessment, video surveillance, privacy protection, image and video analysis, multimedia content search and retrieval, and video transmission over wireless network. He is author or co-author of 3 books, more than 200 research publications and 20 patents issued or pending.



Xin Lu Xin Lu received the B.Sc. and M.Sc. degrees from the Harbin Institute of Technology (HIT), Harbin, China, in 2008 and 2010, respectively, and the Ph.D. degree in computer science from the University of Warwick, Coventry, U.K., in 2013. He is currently a Senior Lecturer with the School of Computer Science and Informatics at De Montfort University (DMU) in Leicester, UK. Before joining DMU, he was a Lecturer (Assistant Professor) in the School of Electronics and Information Engineering at HIT in China. His current research interests in-

clude video coding standards, data compression, deep learning, convolutional neural network, multimedia coding and transmission, and pattern recognition.



Bo Hang Bo Hang received his M.S. degree in computer science from Harbin Engineering University, Harbin, Heilongjiang province, P.R.China, in 2003, and received his Ph.D. degree in Communication and Information System from Wuhan University, Wuhan, Hubei province, P.R.China in 2012. Now he is a professor of Computer School, Hubei University of Arts and Science, Xiangyang, Hubei province, P.R.China. His current research interests include multimedia information processing and computer networks.