Fast Texture Coding Based on Spatial, Temporal and Inter-View Correlations for 3D Video Coding

JIE-RU LIN\textsuperscript{1}, MEI-JUAN CHEN\textsuperscript{1}, (Senior Member, IEEE), YI-SHENG CIOU\textsuperscript{1}, CHIA-HUNG YEH\textsuperscript{2,3}, (Senior Member, IEEE), MIN-HUI LIN\textsuperscript{3}, LIH-JEN KAU\textsuperscript{4}, (Senior Member, IEEE), AND CHUAN-YU CHANG\textsuperscript{5}, (Senior Member, IEEE)

\textsuperscript{1}Department of Electrical Engineering, National Dong Hwa University, Hualien 974301, Taiwan
\textsuperscript{2}Department of Electrical Engineering, National Taiwan Normal University, Taipei 106308, Taiwan
\textsuperscript{3}Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 804201, Taiwan
\textsuperscript{4}Department of Electronic Engineering, National Taipei University of Technology, Taipei 106344, Taiwan
\textsuperscript{5}Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Yunlin 640301, Taiwan

Corresponding author: Chia-Hung Yeh (chyeh@ntnu.edu.tw)

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 105-2221-E-259-016-MY3, Grant MOST 108-2221-E-027-027, Grant MOST 110-2218-E-110-008-, and Grant MOST 110-2218-E-003-004-.

\textbf{ABSTRACT} 3D-HEVC (The 3D Extension of High Efficiency Video Coding), the latest international standard for 3D video coding, supports multiview plus depth 3D video format to enrich multimedia applications. For the texture coding, 3D-HEVC utilizes not only the information of temporal and spatial domains but also that of inter-view domain. However, the time consumption and complexity of 3D-HEVC also increase significantly. In this paper, fast texture coding algorithm for 3D-HEVC is proposed. We individually calculate the Pearson correlation coefficients by the rate-distortion costs (RD-costs) of coding tree units (CTUs) in the temporal, spatial and inter-view domains to analyze the correlations for independent view and dependent view. The proposed coding algorithm is based on the coding information of CTUs with higher correlations. The fast algorithm predicts and dynamically adjusts the depth range of the coding unit (CU). The prediction unit (PU) mode decision is proposed according to the complexity and partition direction of the best PU modes obtained from the highly correlated CTUs. The search range is adaptively adjusted for motion estimation. In addition, the RD-cost threshold is estimated to early terminate the CU split. Experimental results show that the proposed fast texture coding algorithm reduces the texture coding time significantly and outperforms numerous previous works.

\textbf{INDEX TERMS} 3D-HEVC, texture coding, CU early termination, fast PU decision, search range adjustment.

\section{I. INTRODUCTION}
With advanced multimedia technology, video has been widely used in various aspects. Compared with a typical 2D display, people may be more interested in the immersive user experience of a 3D display. In recent years, 3D video applications related to free viewpoint television (FTV) \cite{1,2}, 360° panoramic video, virtual reality (VR) \cite{3} and augmented reality (AR) \cite{4} are becoming widespread. These applications are very popular in commercial advertising, education, multimedia entertainment, etc. In response to the tendency, the Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) establishes the video coding standard for 3D videos, as known as the 3D extension of high efficiency video coding (3D-HEVC) \cite{5,6}.

3D-HEVC evolves from multi-view video coding (MVC) \cite{7} and multi-view plus depth (MVD) \cite{8} coding. Multi-view-HEVC (MV-HEVC) \cite{9} is the extension of HEVC in multi-view structure. Multi-view videos can be demonstrated on FTV or 3D stereo display, and the viewers are able to watch the scenes from different angles or enjoy the 3D experience. MV-HEVC effectively encodes the videos from multiple viewpoints by removing the inter-view redundancy. However, the more views, the more data are required. As a result, the 3D video format evolves to MVD. Depth map, a gray level image, uses greyscale to represent the relative...
distance between the camera and the objects in a scene. Through depth map and depth-image-based rendering (DIBR) [10] technique, numerous virtual views can be synthesized, which greatly reduces the required captured views. Figure 1 depicts the system diagram of 3D-HEVC. At the encoder side, the input data are multi-view color textures with the associated depth maps. After being compressed by 3D-HEVC encoder, the encoded bitstream can be extracted according to the backend application at the decoder side. For instance, we can project multiple virtual views on FTV via DIBR if the bitstream is completely decoded. Also, the bitstream can be partially decoded to perform 3D stereo display or typical 2D display. Compared to HEVC encoder, 3D-HEVC efficiently encodes the MVD videos by not only reducing the spatial and temporal redundancy, but also thoroughly exploring the inter-view (between views) and inter-component (between color texture and depth map) relationships. The previously encoded components or views are all able to be the reference sources of the current coding component.

Coding unit (CU) is the basic element in HEVC and 3D-HEVC. Before starting the coding process, an input frame is divided into multiple non-overlapped coding tree units (CTUs) by the size of $64 \times 64$. A CU can be further split into CUs with various sizes from Depth 0 ($64 \times 64$) to Depth 3 ($8 \times 8$). The CU split is based on quadtree structure, which performs a top-down strategy to split the CU into every possible CU size to find the best combination of the CU partitions with the minimum rate-distortion cost (RD-cost). Besides, each CU contains at least one prediction unit (PU) and one transform unit (TU). PU partition modes deal with different texture variations. PU only allows square partitions ($2N \times 2N$ and $N \times N$) for intra prediction. For inter-frame prediction, in addition to $2N \times 2N$ and $N \times N$, PU also supports symmetric mode partition (SMP; $2N \times N$ and $N \times 2N$) and asymmetric mode partition (AMP; $2N \times nU$, $2N \times nD$, $nL \times 2N$ and $nR \times 2N$). Similar to HEVC, 3D-HEVC also reduces the inter-frame redundancy by performing motion compensated prediction (MCP) on temporal reference frames within a given search range (SR) to explore the best matching blocks. In addition, when coding the dependent views, the independent view has already been encoded and is able to be the reference source. As a result, 3D-HEVC can also implement disparity compensated prediction (DCP) to find the best matching block at the independent view. In summary, MCP refers to the inter-frame prediction by using the already encoded temporal frames of the same view, while DCP indicates the inter-view prediction by using the already encoded frames of the reference views in same access unit.

The fast algorithms to accelerate the coding of 3D-HEVC has been investigated lastingly [11]–[14]. Several mechanisms have been proposed for fast color texture coding. Heo et al. [15] early terminate the construction of the Merge list by checking the most probable candidates and the motion data of temporally collocated PU. Ding et al. [16] determine the range of the CU depth level by analyzing the weighted motion homogeneity of spatial, temporal and inter-view coded tree-blocks. Song and Jia [17] propose an early Merge mode decision for the dependent view by considering the encoded PU modes of inter-view reference CTU and parent CU. Zhang et al. [18] acquire the RD-costs of the spatial,
temporal and inter-view neighboring CTUs. Then, a RD-cost threshold for the early termination of mode decision process is computed by the weighted RD-costs from these neighboring CTUs. Pan et al. [19] adaptively reduce the spatial and temporal MV candidates for the PU with the Merge mode. For the inter mode PU, the combination of spatial and temporal candidates is utilized to early determine the final MV. Zhang et al. [20] use the motion homogeneity model to propose a fast texture video coding method including the fast depth level range determination, early SKIP/Merge mode decision and adaptive motion search range adjustment. Bakkouri and Elyousf [21] use the features based on CU homogeneity as the input for Fuzzy C-Means clustering algorithm to decide the quadtree splitting of CU and choose the thresholds for early termination of CU partition. Zhang et al. [22] propose fast depth level, early SKIP/Merge mode and adaptive early termination based on the correlations of inter-view, spatio-temporal, and texture-depth to decide to skip some treeblocks of texture and depth map at the early stage. This algorithm applies to both texture and depth map, and significantly reduces the encoding time. Avila et al. [23] take advantage of the depth map information to calculate the sample average (SA) and sample gradient (SG) for the decision of background and foreground-ROI. Then, the CU split is early terminated if current CU is in non-ROI region. Bakkouri et al. [24] characterize the homogeneity of a CU by calculating the gradient vectors and the structure tensor. Then, a fast algorithm combining early termination of CU split and mode prediction process is accomplished for both the color texture and depth map. Chen et al. [25] analyze inter-view correlation by observing the best mode distribution for the independent view and dependent view. The early Skip mode decision prejudges whether Skip is the best mode by referring the inter-view neighboring CTUs. Also, the inter SMP and AMP modes are selectively disabled by checking the difference of two adjacent pixels in main diagonal of the CU. Li et al. [26] jointly exploit the priori and posterior probability models to early verify whether Merge mode is the best or not for the dependent view.

Although the fast algorithms for HEVC are getting mature, the time-saving and coding performance are still limited if we just duplicate the strategies from HEVC to 3D-HEVC. The reason is that the properties of the multi-view structure are not fully considered. Moreover, when designing the fast algorithms for the color texture coding of 3D-HEVC, we should be cautious that the coding efficiency should be strictly maintained because the coding results of each view or each component interact with each other. Otherwise, the prediction error may propagate to the depth map, and even the synthesized view. Some works take all the spatial, temporal and inter-view correlations into consideration by using the weighting factors [16], [18], [20], [22]. However, we do not know the appropriate priority of the information. These weighted factors vary from different situations. Sometimes, the critical information may be ignored. Based on the above discussions, this paper considerably deliberates upon spatial, temporal, and inter-view relationships by calculating the Pearson correlation firstly. Secondly, we extract the essential reference information with the higher correlation. Then, the CU depth range prediction and adjustment, fast inter PU mode decision, search range adjustment and CU early termination based on RD-cost estimation are proposed.

II. PROPOSED METHOD
A. ANALYSIS OF SPATIAL, TEMPORAL AND INTER-VIEW CORRELATIONS
1) PEARSON CORRELATION COEFFICIENT
Figure 2 demonstrates the RD-cost distributions for the dependent view at the 6-th frame for Kendo sequence, including the current coding frame, forward-reference frame,
backward-reference frame and inter-view frame. According to Figure 2(a) (b) (c) (d), the RD-cost distributions are quite similar. Moreover, the overlapped RD-cost distributions in Figure 2(e) verify the RD-cost dependencies between the reference frames. As a result, we utilize the RD-costs of the neighboring CTUs to calculate the Pearson correlation coefficients and analyze the spatial, temporal and inter-view relationships.

Eq. (1) formulates the Pearson correlation coefficient ($\rho_{X,Y}$), which is commonly used in statistics to investigate the linear relationship between two sets of variables ($X$ and $Y$). The variables $x_i$ and $y_i$ represent the $i$-th element in the set $X$ and set $Y$, respectively. The variable $n$ indicates the number of the components in the set. The value of $\rho_{X,Y}$ ranges from $-1$ to $1$. The larger the $\rho_{X,Y}$ is, the more relative the two sets.

$$\rho_{X,Y} = \frac{\sum x_i y_i - (\sum x_i \sum y_i) / n}{\sqrt{\sum (x_i)^2 - (\sum x_i)^2 / n}} \sqrt{\sum (y_i)^2 - (\sum y_i)^2 / n}$$ (1)

When encoding independent view, the reference sources are only spatial and temporal domains; when encoding dependent view, inter-view domain is also available. Thus, the analysis of the independent view and dependent view are performed individually.

2) INDEPENDENT VIEW

We investigate the spatial and temporal relationships for the analysis of the independent view coding. In terms of spatial relationship, we analyze the correlations between the current coding CTU and the left (L), upper (U), upper-left (UL) and upper-right (UR) CTUs. In terms of the temporal relationship, the correlations between the current coding CTU and the forward-reference collocated (C0) CTU as well as the backward-reference collocated (C1) CTU are inspected. Figure 3 specifies all the reference CTUs for the spatial and temporal correlation analyses of the independent view.

![Figure 3. Reference CTUs for the spatial and temporal correlation analyses of the independent view.](image-url)

We collect the RD-costs ($RD$s) of these reference CTUs and calculate the Pearson correlation coefficient by Eq. (1) for each reference neighboring location. Eq. (2) clusters the sets for the calculation of the spatial correlation ($X_{I,S}$ and $Y_{I,SA}$), and Eq. (3) groups the sets for the calculation of the temporal correlation ($X_{I,T}$ and $Y_{I,TB}$). Besides, the RD-cost of the current coding CTU is unknown, so $RD_{I,C,avg}$ is added into Eqs. (2) and (3) to augment the sample sets. $RD_{I,C,avg}$ is obtained by averaging the RD-costs of the forward-reference ($RD_{I,C0}$) and the backward-reference ($RD_{I,C1}$) CTUs. Finally, there are four spatial correlation coefficients and two temporal correlation coefficients as in the coefficient set ($\rho_{X,Y}^{I}$), which is formulated in Eq. (4).

$$X_{I,S} = \{RD_{I,C,avg}, RD_{I,C0}, RD_{I,C1}\}$$
$$Y_{I,SA} = \{RD_{I,a}, RD_{I,0,a}, RD_{I,a}|\alpha \epsilon \{L, U, UL, UR\}\}$$ (2)

$$X_{I,T} = \{RD_{I,C,avg}, RD_{I,L}, RD_{I,U}, RD_{I,UL}, RD_{I,UR}\}$$
$$Y_{I,TB} = \{RD_{I,C,\beta}, RD_{I,L,\beta}, RD_{I,U,\beta}, RD_{I,UL,\beta}, RD_{I,UR,\beta} \mid \beta \epsilon \{0, 1\}\}$$ (3)

$$\rho_{X,Y}^{I} = \{\rho_{X,I,S,Y,I,SL}, \rho_{X,I,S,Y,FU}, \rho_{X,I,S,Y,UL}, \rho_{X,I,S,Y,UR}, \rho_{X,I,T,Y,LT}, \rho_{X,I,T,Y,UT}\}$$ (4)

3) DEPENDENT VIEW

At the same access unit, the independent view is already coded before dependent view coding, so the inter-view relationship can be also taken into consideration. For the coding of the dependent view, the correlation analyses of the spatial and temporal relationships are inherited from the way of the independent view. The correlations between the current coding CTU and the left (L), upper (U), upper-left (UL), upper-right (UR), forward-reference collocated (C0), and backward-reference collocated (C1) CTUs are also explored by acquiring the RD-costs ($RD$s) of these reference CTUs and calculating the Pearson correlation coefficients by Eq. (1). The sample sets for the calculation of the spatial correlation ($X_{D,S}$ and $Y_{D,SA}$) are grouped in Eq. (5), while the sample sets for the calculation of temporal correlation ($X_{D,T}$ and $Y_{D,TB}$) are clustered in Eq. (6).

$$X_{D,S} = \{RD_{D,C,avg}, RD_{D,C0}, RD_{D,C1}, RD_{IV,C}\}$$
$$Y_{D,SA} = \{RD_{D,a}, RD_{D,0,a}, RD_{D,a} \mid \alpha \epsilon \{L, U, UL, UR\}\}$$ (5)

$$X_{D,T} = \{RD_{D,C,avg}, RD_{D,L}, RD_{D,U}, RD_{D,UL}, RD_{D,UR}\}$$
$$Y_{D,TB} = \{RD_{D,C,\beta}, RD_{D,L,\beta}, RD_{D,U,\beta}, RD_{D,UL,\beta}, RD_{D,UR,\beta} \mid \beta \epsilon \{0, 1\}\}$$ (6)

There are still slight differences from independent view. In addition to the averaged RD-cost ($RD_{D,C,avg}$) from the temporal collocated CTUs ($RD_{D,C0}$ and $RD_{D,C1}$), we shift the current coding CTU from the dependent view to the independent view by using the depth oriented neighboring block based disparity vector (DoNBDV) to acquire the inter-view corresponding CTU (IV_C) and the inter-view corresponding spatial neighboring CTUs. Then, the RD-costs of these inter-view corresponding CTUs are added into the sets for the calculation of the spatial correlation as listed in Eq. (5). All the reference CTUs for the spatial and temporal correlation analyses of the dependent view are specified in Figure 4.

Regarding the inter-view relationship, as shown in Figure 5, we acquire the collocated (C), left (L), upper (U),
right (R) and bottom (B) reference CTUs in the forward (list_0) and backward (list_1) directions. In addition, we shift these temporal neighboring CTUs from the dependent view to the independent view by DoNBDV to obtain the inter-view corresponding CTUs. Finally, the RD-costs of these inter-view reference corresponding CTUs are collected and the calculation of the inter-view correlation is also conducted by Eq. (1), which the sample sets \(X_{D_{STIV}}\) and \(Y_{D_{STIV}}\) are enumerated in Eq. (7). In summary, there are four spatial correlation coefficients, two temporal correlation coefficients, and one inter-view correlation coefficients as in the coefficient set \(\rho_{D_{STIV}}\), which is formulated in Eq. (8).

\[
X_{D_{IV}} = \{RD_{D_{C0}}, RD_{D_{L0}}, RD_{D_{U0}}, RD_{D_{R0}}, RD_{D_{C1}}, RD_{D_{L1}}, RD_{D_{U1}}, RD_{D_{R1}}, RD_{D_{B1}}\}
\]

\[
y_{D_{IV}} = \{RD_{IV_{C0}}, RD_{IV_{L0}}, RD_{IV_{U0}}, RD_{IV_{R0}}, RD_{IV_{C1}}, RD_{IV_{L1}}, RD_{IV_{U1}}, RD_{IV_{R1}}\}
\]

\[
\rho^D_{X,Y} = \left\{ \rho^D_{X_{DS}, Y_{DS}}, \rho^D_{X_{DS}, Y_{DSU}}, \rho^D_{X_{DS}, Y_{DSUR}}, \rho^D_{X_{DS}, Y_{DSUL}} \right\}
\]

4) EXTRACTION OF THE REFERENCE INFORMATION

In this paper, we want to utilize the encoded information from left, upper, upper-left, upper-right, forward collocated, backward collocated and inter-view corresponding CTUs to accelerate the entire coding process either in CU or PU level. Instead of coarsely adopting all the information from these reference CTUs, we only choose the critical information from the CTUs with the higher correlations based on the correlation analysis. The inaccurate or unrelated reference information can be dropped out to avoid the decision error when performing the proposed fast algorithm.

After the calculations of the Pearson correlation coefficients, we sort the elements in the coefficient set \(\rho^I_{ST}\) for the independent view and \(\rho^D_{X,Y}\) for the dependent view. The encoded information is extracted from the top M CTUs with the highest correlations for the fundamental of the following proposed algorithms. To decide the appropriate quantity of the reference CTUs, we investigate the correlation coefficients by encoding the test sequences (Kendo, Balloons, Newspaper, Poznan_Street, Poznan_Hall2, Shark, Undo_Dancer, GT_Fly). Table 1 and Table 2 tabulate the correlation rankings of the independent view and dependent view, respectively. We can notice that the third highest correlations of both independent view and dependent view have been downed to around 0.6. In addition, there is an obvious
gap between the third and the fourth highest correlations and the correlations after the third highest correlations are relative lower. Thus, $M$ is set as 3 in the experiments to balance the time-saving performance and coding efficiency.

### B. CU DEPTH RANGE PREDICTION AND DYNAMIC ADJUSTMENT

In this section, the proposed CU depth range prediction and dynamic adjustment are described in detail. To explain clearly, the related parameters are defined in Table 3.

#### 1) CU DEPTH RANGE PREDICTION

Based on the correlation analysis, we predict the CU depth range from the top $M$ reference CTUs with the highest correlations as shown in Eq. (9). The lower bound ($Depth_{\min}$) and the upper bound ($Depth_{\max}$) of the predicted CU depth range are determined by the minimum and the maximum CU depths among these reference CTUs, respectively. Compared to the ordinary criterion which directly predicts the CU depth from all the reference CTUs in spatial, temporal and inter-view domains, the proposed CU depth range prediction has selection mechanism on the fundamental of correlation analysis.

\[
Depth_{\min} = \min \{Depth_m\}, \quad m \in \{1, 2, \ldots, M\} \quad (9)
\]

\[
Depth_{\max} = \max \{Depth_m\}, \quad m \in \{1, 2, \ldots, M\}
\]

In addition, Merge/Skip or 2Nx2N mode is often selected as the best mode for the PU. If we immediately bypass the CU depth which is out of the predicted range, it may result in significant impact on coding efficiency. Therefore, if the current CU depth ($Depth_{\text{cur}}$) is not within the predicted CU depth range, we reserve the prediction of Merge/Skip and 2Nx2N modes and omit the other PU modes to reduce the coding time. Besides, intra prediction is unchanged.

#### 2) DYNAMIC ADJUSTMENT FOR CU DEPTH RANGE

Although the proposed CU depth range prediction has thoroughly investigated the spatial, temporal and inter-view correlations, the video content, parameters and frame types between the current coding CTU and the reference CTUs are still slightly different. The possibility of the misjudgment cannot be completely avoided. To avoid the prediction error, we make use of the coding characteristics of the current coding CTU and dynamically adjust the predicted CU depth range. Three kinds of strategies are conceived to fine-tune the CU depth range as follows.

**Algorithm Dynamic Adjustment for CU Depth Range**

- **if** $Depth_{\text{range}} \leq 2$ and $Depth_{\min} \neq 0$ then
  \[Depth_{\min} \leftarrow Depth_{\min} - 1.\]
  **end if**
- **if** Inter_SMP $\notin PU_{\text{ref}}$ and Inter_AMP $\notin PU_{\text{ref}}$ and $PU_{\text{best}} = \text{Merge/ Skip}$ and $RD_{\text{normalized}} < RD_{\text{ref, avg}}$ then
  \[Depth_{\min} \leftarrow Depth_{\min} - 1.\]
  **end if**
- **if** (Inter_SMP $\in PU_{\text{ref}}$ or Inter_AMP $\in PU_{\text{ref}}$) and $PU_{\text{best}} \neq \text{Merge/ Skip and } Depth_{\text{range}} \leq 1$ and $RD_{\text{normalized}} \geq RD_{\text{ref, avg}}$ then
  \[Depth_{\max} \leftarrow Depth_{\max} + 1.\]
  **end if**
The $RD_{\text{normalized}}$ is calculated by Eq. (10)

$$RD_{\text{normalized}} = 4^{\text{Depth}_{\text{cur}}} \times RD_{\text{cur}}$$  (10)

The $\text{Depth}_{\text{min}}$ is reduced by 1 to additionally include the larger size CU for smaller predicted depth range ($\text{Depth}_{\text{range}}$) by the first condition and possibly smother CU content by the second condition. The $\text{Depth}_{\text{max}}$ is increased by 1 to further check the smaller size CU for possibly complex CU partition. The judgements of the latter two conditions mainly come from the information of reference PUs and RD-Costs.

C. FAST INTER PU MODE DECISION

The best PU mode at current CU depth is usually similar to the PU modes in reference CTUs. As a result, the proposed fast inter PU mode decision anticipates the potential PU partitions in advance to accelerate the prediction process. Large proportion of the prediction time can be decreased if we directly restrict the PU partitions to the PU modes which have been already chosen in the reference CTUs. However, the coding efficiency could be dramatically degraded.

During the coding of the independent view, the accessible reference information is comparatively less (no inter-view information). Moreover, the prediction error will also propagate to dependent view through inter-view dependency. Referring to the hierarchical coding order of the random-access prediction structure, the frame distances between either forward-reference or backward-reference frames at the lower hierarchical layers are larger, so the prediction errors are accordingly more evident. Also, if lots of prediction errors are generated in the lower hierarchical layers, these errors are very likely to drift to the higher hierarchical layers. Hence, a strict mode decision strategy is necessary to the coding of the independent view or lower hierarchical layers. In contrast, a powerful mode decision strategy is feasible during the coding of the dependent view coding or the higher hierarchical layers.

Several parameters used in this section are also defined in Table 3. There are two kinds of decision strategies in the proposed fast inter PU mode decision. The first one is based on mode complexity, which is for the independent view coding or the dependent view coding at lower hierarchical layer. The set composed of the PU modes which have been already chosen in the reference CTUs is defined as $PU_{\text{ref}}$. If $PU_{\text{ref}}$ does not include the complicated asymmetric inter mode partitions ($\text{Inter}_\text{AMP}$), the prediction of $\text{Inter}_\text{AMP}$ mode at current CU depth is disabled. If both the asymmetric inter mode partitions ($\text{Inter}_\text{AMP}$) and symmetric inter mode partitions ($\text{Inter}_\text{SMP}$) do not belong to $PU_{\text{ref}}$, the prediction of $\text{Inter}_\text{AMP}$ and $\text{Inter}_\text{SMP}$ mode at current CU depth is disabled simultaneously.

The secondary strategy is based on partition directions, which is for the dependent view coding at higher hierarchical layer. If $PU_{\text{ref}}$ is not composed of any horizontal inter partition mode ($\text{Inter}_\text{Hor}$), the prediction of $\text{Inter}_\text{Hor}$ at current CU depth is inactivated. Likewise, the vertical inter partition mode ($\text{Inter}_\text{Ver}$) is bypassed at current CU depth if no element in $PU_{\text{ref}}$ belongs to $\text{Inter}_\text{Ver}$. This method restricts the direction of PU partitions in a straightforward manner, which can reduce much coding time if the reference information is correct enough.

If any neighboring PU selects intra mode as the best mode, it means that not all the reference information refers to the properties of inter prediction, which might cause the decision error of the proposed criteria for inter partitions. To avoid this special case, we execute an intra-boundary check on the border pixels of the $M$ reference CTUs. The checked pixels are illustrated in Figure 6.

![Intra-boundary check of the reference CTUs.](image)

For the left and upper reference CTUs, the checked pixels are the border pixels of the adjacent encoded PU (blue-highlighted). For the upper-left and upper-right reference CTUs, the checked pixels are the connecting corner pixel (yellow-highlighted). For the forward collocated, backward collocated and inter-view corresponding CTUs, the checked pixels are the four corner pixels (pink-highlighted). $\text{Intra}_\text{check}_\text{PU}$ specifies how many checked pixels are encoded as intra modes. The proposed fast inter PU mode decision is operated only if $\text{Intra}_\text{check}_\text{PU}$ is zero (no intra mode is selected as the best mode among the neighboring encoded PUs). The complete flowchart of the proposed fast intra mode decision is demonstrated in Figure 7.

D. SEARCH RANGE ADJUSTMENT

The magnitude of the search range (SR) affects not only the precision but also the time-consumption when executing motion estimation. The most common way is to reduce the search range by a pre-defined scaling factor according to some parameters, such as block complexity. A more flexible method is conceived in the proposed search range adjustment.

We collect the motion vectors (MVs) of all PUs from the reference CTUs. Then, the search range is modified as the maximum MV length among the MVs from the reference CTUs ($\| MV_{\text{ref}_{\text{max}}} \|_2$). The calculation of $\| MV_{\text{ref}_{\text{max}}} \|_2$ is formulated in Eq. (11). $\| MV_{\text{ref}} \|_2$ represents the maximum MV length among the MVs from the $m$-th reference CTU. $M$ denotes the number of the reference CTUs with the highest correlations, which is the same parameter in
previous sections.

$\left\| MV_{ref,max} \right\|_2 = \max_{m \in \{1, 2, \ldots, M\}} \left\| MV_{ref,m} \right\|_2$, 

Some special situations are worth discussing. If the AMVP vector of the current PU is $(0,0)$, there are two meanings. The first case is that the best MV candidate of AMVP is $(0,0)$ exactly. The second case is that the reference location of AMVP is encoded as intra mode and the MV is not available. Especially for the second case, the start point of the motion estimation might be inaccurate, so the search range should be large enough. In addition, if the length of the AMVP vector is larger than $\left\| MV_{AMVP} \right\|_2$, the ground-truth MV of the current PU is probably outside the modified search range. The best matching block with the minimum motion cost is then inaccessible if we greatly reduce the search range under the cases mentioned above. Therefore, some restrictions are bounded to the proposed search range adjustment. If the length of the AMVP vector $\left\| MV_{AMVP} \right\|_2$ is larger than the modified search range $\left\| MV_{ref,max} \right\|_2$ or the number of intra check points (Figure 6) at the AMVP reference candidates $\text{Intra}_{\text{count}}$ is not zero, the search range remains unchanged. $\left\| MV_{AMVP} \right\|_2$ is formulated in Eq. (12),

$E. \text{CU EARLY TERMINATION BASED ON RD-COST ESTIMATION}$

The proposed method estimates a threshold for the best RD-cost of the current coding CU to early terminate the CU split. Several related parameters are defined in Table 4. The threshold for the CU early termination ($TH_{ET}$) can be determined and adjusted based on the RD-cost distributions of the reference CTUs.

In this paper, the RD-costs are utilized to evaluate the spatial, temporal and inter-view correlations. The RD-cost distributions seem similar if the CTUs are highly correlated. We use this property by estimating the best RD-cost of the CU for the current CU depth ($RD_{m,Depth_{cur,predict}}$) from the $m$-th reference CTU for CU early termination. The procedures are illustrated in Figure 8 and described below.

Step 1: Acquire the minimum and maximum RD-costs ($\min RD_{m,D}$ and $\max RD_{m,D}$) among all the CTUs with a maximum CU depth $D$ in the frame of the $m$-th reference CTU. In addition, $m \in \{1, 2, \ldots, M\}$ and $D \in \{0, 1, 2, 3\}$.

Step 2: Obtain the maximum CU depth ($Depth_{ref}$) and the RD-cost ($RD_{m}$) of the $m$-th reference CTU.

Step 3: Predict the RD-cost of the CTU for the current CU depth ($RD_{m,Depth_{cur,predict}}$) based on the RD-cost of $m$-th reference CTU by Eq. (13), which is a linear scaler converting $RD_{m}$ at $Depth_{ref}$ to the predicted RD-cost ($RD_{m,Depth_{cur,predict}}$) at $Depth_{cur}$.

$RD_{m,Depth_{cur,predict}} = \left( \frac{RD_{m} - \min RD_{m,Depth_{ref}}}{\max RD_{m,Depth_{cur}} - \min RD_{m,Depth_{cur}}} \right) \times (\max RD_{m,Depth_{cur}} - \min RD_{m,Depth_{cur}}) + \min RD_{m,Depth_{cur}}, \quad m \in \{1, 2, \ldots, M\}$

1) DETERMINATION OF THE LARGER EXTREME FOR $RD_{m}$

If the RD-cost of $m$-th reference CTU ($RD_{m}$) is an inordinate larger value, we are not sure that $RD_{m}$ is the RD-cost with either good or bad prediction. If $RD_{m}$ belongs to the latter one, the estimated best RD-cost by our method may be valueless.

In probability theory, for any probability distribution, the Chebyshev Inequality [27] ensures the percentage of the random variables which is apart from the mean value by a given scaling is not larger than a certain fraction. To avoid the error decision, we adopt the One-side Chebyshev Inequality [28] as a discriminator to distinguish the inordinate RD-cost, which is expressed in Eq. (14). In Eq. (14), One-side Chebyshev Inequality [28] guarantees the probability that $RD_{m}$ is larger than or equal to the mean RD-cost ($\mu_{m,Depth_{cur}}$) added by $k_1$ times of the standard deviation ($\sigma_{m,Depth_{cur}}$) is smaller than or equal to $1/(1+k_1^2)$. As the blue-painted region in the Figure 9, if $RD_{m}$ is larger than or equal to $\mu_{m,Depth_{cur}} + k_1 \times \sigma_{m,Depth_{cur}}$, 

\[
\left\| MV_{AMVP} \right\|_2 = \sqrt{MV_{x,AMVP}^2 + MV_{y,AMVP}^2} \quad (12)
\]
we recognize RD\textsubscript{m} as a larger extreme. In our design, we set probability (1/1+k\textsubscript{1}) as 30%, so that k\textsubscript{1} is 1.528 approximately. If RD\textsubscript{m} is a larger extreme, the CU early termination based on RD-cost estimation is disabled.

\[ P(\text{RD}_m \geq \mu_{\text{Depth}_\text{ref}} + k_1 \times \sigma_{\text{Depth}_\text{ref}}) \leq \frac{1}{1+k_1^2} \]  \hspace{1cm} (14)

2) DISCRIMINANION OF THE SMALLER EXTREME FOR RD\textsubscript{m,Depth\textsubscript{cur,predict}}

If the RD-cost is comparably low, it usually indicates that the prediction result is quite accurate. Therefore, the predicted RD-cost does not need adjustment. After the predicted best RD-cost (RD\textsubscript{m,Depth\textsubscript{cur,predict}}) is obtained, we combine the One-side Chebyshev Inequality to determine whether the RD-cost is a smaller extreme, which is expressed in Eq. (15). As the green-painted region in the Figure 10, if RD\textsubscript{m,Depth\textsubscript{cur,predict}} is smaller than or equal to \( \mu_{\text{Depth}_\text{cur}} - k_2 \times \sigma_{\text{Depth}_\text{cur}} \), the RD-costs located in this region are with the relatively accurate prediction results and we recognize RD\textsubscript{m,Depth\textsubscript{cur,predict}} as a smaller extreme. In this case, no additional adjustment is required for RD\textsubscript{m,Depth\textsubscript{cur,predict}}. The predicted RD-cost from the m\textsubscript{th} reference CTU for CU early termination is determined by RD\textsubscript{m,ET}, which is formulated in Eq. (16). In our experiment, the probability of 1/1+k\textsubscript{2} is set as 5%, so that k\textsubscript{2} is 4.359 approximately.

\[ P\left(\text{RD}_m,\text{Depth}_\text{cur,predict} \leq \mu_{\text{Depth}_\text{cur}} - k_2 \times \sigma_{\text{Depth}_\text{cur}}\right) \leq \frac{1}{1+k_2^2} \] \hspace{1cm} (15)

\[ \text{RD}_{m,ET} = \text{RD}_{m,\text{Depth}_\text{cur,predict}}, \text{if } \text{RD}_{m,\text{Depth}_\text{cur,predict}} \leq \mu_{\text{Depth}_\text{cur}} - k_2 \times \sigma_{\text{Depth}_\text{cur}} \] \hspace{1cm} (16)

3) ADAPTIVE ADJUSTMENT FOR RD\textsubscript{m,Depth\textsubscript{cur,predict}}

In order to make the predicted best RD-cost more rigorous, an amendment is operated after obtaining RD\textsubscript{m,Depth\textsubscript{cur,predict}}. Instead of directly setting an adjustment parameter, we perform an adaptive adjustment according to the distribution of RD\textsubscript{m,Depth\textsubscript{cur,predict}} and calculate the RD-cost (RD\textsubscript{m,ET}) for CU early termination. If RD\textsubscript{m} is not a larger extreme and RD\textsubscript{m,Depth\textsubscript{cur,predict}} is not a smaller extreme, we perform an adaptive adjustment on RD\textsubscript{m,Depth\textsubscript{cur,predict}} to make the threshold for CU early termination stricter. For a general normal distribution, if RD\textsubscript{m,Depth\textsubscript{cur,predict}} is closer to the mean value (\( \mu_{\text{Depth}_\text{cur}} \)), it indicates that RD\textsubscript{m,Depth\textsubscript{cur,predict}} is probably not an outlier, which is more reliable. Only slight adjustment is required. On the contrary, substantially adjustment is necessary if RD\textsubscript{m,Depth\textsubscript{cur,predict}} is far away from the mean value (\( \mu_{\text{Depth}_\text{cur}} \)).

As the yellow-painted region in Figure 11, when RD\textsubscript{m,Depth\textsubscript{cur,predict}} is larger than \( \mu_{\text{Depth}_\text{cur}} \), the estimated RD-cost for CU early termination (RD\textsubscript{m,ET}) is obtained from adjusting RD\textsubscript{m,Depth\textsubscript{cur,predict}} by Eq. (17). Likewise, as the yellow-painted region in Figure 12, RD\textsubscript{m,ET} is calculated from adjusting RD\textsubscript{m,Depth\textsubscript{cur,predict}} by Eq. (18) when RD\textsubscript{m,Depth\textsubscript{cur,predict}} is smaller than \( \mu_{\text{Depth}_\text{cur}} \). Eqs. (17) and (18) linearly scale RD\textsubscript{m,Depth\textsubscript{cur,predict}} by a fraction
relative to the distance away from the mean value ($\mu_{m, \text{Depth}_\text{cur}}$). Besides, if $RD_{m, \text{Depth}_\text{cur}, \text{predict}}$ is equal to $\mu_{m, \text{Depth}_\text{cur}}$, no adjustment is performed.

$$RD_{m, \text{Depth}_\text{cur}, \text{ET}} = RD_{m, \text{Depth}_\text{cur}, \text{predict}} \times \left( \frac{\max RD_{m, \text{Depth}_\text{cur}, \text{predict}} - RD_{m, \text{Depth}_\text{cur}, \text{predict}}}{\max RD_{m, \text{Depth}_\text{cur}} - \mu_{m, \text{Depth}_\text{cur}}} \right),$$

if $RD_{m, \text{Depth}_\text{cur}, \text{predict}} > \mu_{m, \text{Depth}_\text{cur}}$ (17)

$$RD_{m, \text{ET}} = RD_{m, \text{Depth}_\text{cur}, \text{predict}} \times \left( \frac{RD_{m, \text{Depth}_\text{cur}, \text{predict}} - \min RD_{m, \text{Depth}_\text{cur}, \text{predict}}}{\mu_{m, \text{Depth}_\text{cur}} - \min RD_{m, \text{Depth}_\text{cur}} - \mu_{m, \text{Depth}_\text{cur}}} \right),$$

if $RD_{m, \text{Depth}_\text{cur}, \text{predict}} < \mu_{m, \text{Depth}_\text{cur}}$ (18)

4) FINAL THRESHOLD FOR CU EARLY TERMINATION ($TH_{\text{ET}}$)

Lastly, we choose the minimum $RD_{m, \text{ET}}$ as the final threshold for CU early termination ($TH_{\text{ET}}$) as Eq. (19). After finishing all the prediction at the current CU depth, if the RD-cost of the best mode is smaller than the estimated RD-cost threshold ($TH_{\text{ET}}$), it implies that the prediction result at the current CU depth is very precise. Under this situation, we early terminate the following CU split to diminish the coding time. Besides, if $PU_{\text{ref}}$ are all Merge/Skip, the current best PU mode ($PU_{\text{best}}$) is Merge/Skip, and $RD_{\text{normalized}}$ is smaller than the average of the RD-costs of the reference CTUs ($RD_{\text{ref,avg}}$), the CU split is also early terminated.

$$TH_{\text{ET}} = \min \{RD_{m, \text{ET}}\}, \ m \in \{1, 2, \ldots, M\} \quad (19)$$

F. OVERALL ALGORITHM

The flowchart of the overall fast algorithm is demonstrated in Figure 13, which includes the proposed CU depth range prediction and dynamic adjustment, fast inter PU mode decision, search range adjustment and CU early termination based on RD-cost estimation.

III. EXPERIMENTAL RESULTS

The proposed fast texture coding algorithm is implemented in the 3D-HEVC reference software version 16.0 (HTM-16.0) [29]. Three-view case with associated depth maps is encoded under random-access configuration. Related settings for the experimental environment completely follow the Common Test Condition (CTC) [30] from JCT3V, which are listed in Table 5.

The test benchmark sequences are tabularized in Table 6. There are two kinds of resolutions ($1024 \times 768$ and $1920 \times 1088$). These test sequences can be classified into two classes, including the nature-scene videos (Kendo, Balloons, Newspaper, Poznan_Street, Poznan_Hall2) and the artificial animations (Shark, Undo_Dancer, GT_Fly). The evaluation of
The time-saving performance for the color texture coding time ($T_{\text{texture}}$) is formulated in Eq. (20). Moreover, the coding efficiency is assessed by the Bjøntegaard Delta Bitrate (BD-BR) [31] and the Bjøntegaard Delta Peak Signal-to-Noise Ratio (BD-PSNR) [32]. In addition, the main purpose of the multi-view plus depth format is synthesizing the virtual views, so the coding efficiency of the synthesized view is also necessary to be evaluated. The quality of the synthesized view is evaluated by averaging the Peak Signal-to-Noise Ratio (PSNR) of the six virtual views. Then, the BD-BR of the synthesized view is calculated by the total bitrate and the PSNR of the synthesized view.

$$TS_{\text{texture}}(\%) = \frac{Time_{\text{HTM, texture}} - Time_{\text{proposed, texture}}}{Time_{\text{HTM, texture}}} \times 100\%$$

Table 7 shows the experimental results of the proposed algorithm. V/V, V/T and S/T represent video texture PSNR/video texture bitrate, video texture PSNR/total bitrate, and synthesis view PSNR/total bitrate, respectively. The coding efficiency is well-preserved and the average BD-BR and the average BD-PSNR are only 0.68% and −0.021dB with the 40.75% time-saving for texture coding.

The proposed approach provides more time-saving at the sequences with the larger resolution (Poznan_Street, Poznan_Hall2, Undo_Dancer, GT_Fly). The reason is that the video contents in the sequences with the larger resolution are usually smoother than the sequences with the smaller resolution, and the CTUs with smooth content tend to be Depth 0. Therefore, the CTUs of the sequences with larger resolution (Poznan_Street, Poznan_Hall2, Undo_Dancer, GT_Fly) are more likely to be early terminated, which contribute to more coding time reduction.

Studies [19], [20] provide good time-saving performance for 3D-HEVC texture coding. We make a performance comparison of BD-BR(V/V) and texture coding time between our method with the references [19], [20] in Table 8. The average time-saving of the proposed algorithm in texture coding reach to 40.75%, which is apparently better than 33.03% in [19]. Although the time-saving of texture coding of [20] is larger than that of proposed method, regarding the coding efficiency, the average BD-BR(V/V) of the proposed fast texture coding is superior to [20] by 0.51%. Our method provides much better BD-BR than [20] for Newspaper, Shark, Undo_Dancer and GT_Fly sequences. The performance comparison for the BD-BR(V/T) and $TS_{\text{texture}}$ between the proposed algorithm and [21] is tabulated in Table 9. The proposed algorithm can speed up 40.75% of the texture coding time, which obviously outperforms 29.12% (dependent view) of [21] by 11.63%.
which makes a significant progress. Furthermore, the average BD-BR(V/T) of the proposed approach is only 0.74%.

In addition, Figure 14 demonstrates the comparison of subjective quality for the synthesized view of Newspaper sequence for the proposed algorithm and HTM-16.0. The visual quality looks almost the same. Even in the partial enlargements of the complex region and object boundary, the perceptual quality is almost lossless and the time-saving of the proposed fast algorithm achieves to 34.296%. Figure 15 depicts the RD performance for the synthesized view of Newspaper sequence compared to HTM-16.0. The two RD-curves are very close, which indicates the coding efficiency is well-maintained and similar to the original encoder.
TABLE 8. Performance comparison of our method with [19] and [20].

| Sequences   | [19] BD-RR T Sü texture (%) | [20] BD-RR T Sü texture (%) | Proposed BD-RR T Sü texture (%) |
|-------------|------------------------------|-----------------------------|---------------------------------|
| Kendo       | 0.39                         | 35.25                       | 0.98                            |
| Balloons    | 0.46                         | 32.67                       | 0.91                            |
| Newspaper   | 0.51                         | 31.13                       | 1.45                            |
| Poznan Street | 0.48                        | 34.83                       | 0.57                            |
| Poznan Hall2 | 0.53                        | 36.55                       | 0.38                            |
| Shark       | 0.44                         | 30.26                       | 2.43                            |
| Undo Dancer | 0.38                         | 32.71                       | 2.29                            |
| GT_Fly      | 0.58                         | 30.87                       | 1.16                            |
| **Average** | **0.47**                     | **33.03**                   | **1.27**                        |

TABLE 9. Performance comparison of our method with [21].

| Sequences   | [21] BD-RR T Sü texture,dep (%) | Proposed BD-RR T Sü texture (%) |
|-------------|---------------------------------|---------------------------------|
| Kendo       | 0.68                            | 30.16                           |
| Balloons    | 0.90                            | 30.58                           |
| Newspaper   | 0.94                            | 27.14                           |
| Poznan Street | 0.46                         | 25.57                           |
| Poznan Hall2 | 0.01                        | 28.44                           |
| Shark       | 0.17                            | 33.23                           |
| Undo Dancer | 0.46                            | 26.84                           |
| GT_Fly      | 0.78                            | 30.99                           |
| **Average** | **0.55**                        | **29.12**                       |

IV. CONCLUSION

This paper proposes fast texture coding algorithm for 3D-HEVC. We considerably investigate the spatial, temporal and inter-view relationships by correlation analysis. The essential and critical encoded information from the neighboring CTUs with highest correlations are utilized for the proposed strategies. The proposed CU depth range prediction estimates the probable CU depth interval for the current CTU and bypasses most of the encoding procedures at the unimportant CU depths. Two kinds of fast inter PU decision are conceived, which quickly determine the PU partition types by the complexity and direction of the encoded reference CTUs. The search range of the motion estimation is punctiliously performed with the proposed intra boundary check. A CU early termination criterion is also proposed to avoid the exhausting CU split process. The RD-cost of the current coding CTU is predicted by linear scaling from the RD-costs of the reference CTUs. We analyze the RD-cost distribution of the reference frames and consider various situations based on the one-sided Chebyshev inequality. Experimental results show that we have significantly reduced the time of texture inter coding. Our results outperform the previous works with competitive compromise between computational complexity and RD performance. Subjective comparison verifies that the perceptual quality is well-maintained.

REFERENCES

[1] D. Mielo, O. Stankiewicz, and M. Domanski, “Depth map estimation for free-viewpoint television and virtual navigation,” IEEE Access, vol. 8, pp. 5760–5776, 2020.
[2] A. Smolnic, K. Mueller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and T. Wiegand, “3D video and free viewpoint video–technologies, applications and MPEG standards,” in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2006, pp. 2161–2164.
[3] H. Cao, X. Shang, and H. Qi, “Virtual reality image technology assists training optimization of motion micro-time scale,” IEEE Access, vol. 8, pp. 123215–123227, 2020.
[4] Q. Cheng, S. Zhang, S. Bo, D. Chen, and H. Zhang, “Augmented reality dynamic image recognition technology based on deep learning algorithm,” IEEE Access, vol. 8, pp. 137370–137384, 2020.
[5] G. Tech, Y. Chen, K. Muller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, “Overview of the multiview and 3D extensions of high efficiency video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 1, pp. 35–49, Jan. 2016.
[6] Y. Chen, G. Tech, K. Wegner, and S. Yea, Test Model 11 of 3D-HEVC and MV-HEVC, Document JCT3V-K1003, Feb. 2015.
[7] P. Merkle, A. Smolnic, K. Muller, and T. Wiegand, “Efficient prediction structures for multiview video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 11, pp. 1461–1473, Nov. 2007.
[8] P. Merkle, A. Smolnic, K. Muller, and T. Wiegand, “Multi-view video plus depth representation and coding,” in Proc. IEEE Int. Conf. Image Process., Sep. 2007, pp. I-201–I-204.
[9] M. M. Hannuksela, Y. Yan, X. Huang, and H. Li, “Overview of the multiview high efficiency video coding (MV-HEVC) standard,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2015, pp. 2154–2158.
[10] C. Fehn, “Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV,” Proc. SPIE, p. 5291, pp. 93–104, Jan. 2004.
[11] Q. Zhang, R. Jing, B. Wang, P. Cui, C. Zhou, T. Yan, S. Kanyukt, and W. Si, “Fast mode decision based on gradient information in 3D-HEVC,” IEEE Access, vol. 7, pp. 135448–135456, 2019.
[12] C.-H. Fu, H. Chen, Y.-L. Chan, S.-H. Tsang, H. Hong, and X. Zhu, “Fast depth intra coding based on decision tree in 3D-HEVC,” IEEE Access, vol. 7, pp. 173138–173147, 2019.
[13] H. Hamout and A. Elyousfi, “Fast depth map intra coding for 3D video compression-based tensor feature extraction and data analysis,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 7, pp. 1933–1945, Jul. 2020.
[14] J.-R. Lin, M.-J. Chen, C.-H. Yeh, Y.-C. Chen, L.-J. Kau, C.-Y. Chang, and M.-H. Lin, “Visual perception based algorithm for fast depth intra coding of 3D-HEVC,” IEEE Trans. Multimedia, early access, Mar. 31, 2021, doi: 10.1109/TMM.2021.3070106.
[15] Y. S. Heo, G. Bang, and G. H. Park, “Adaptive merge list construction for 3D-HEVC fast encoder,” Electron. Lett., vol. 52, no. 8, pp. 604–605, Apr. 2016.
[16] H. Ding, W. An, T. Yan, and Q. Zhang, “Fast depth level range determination algorithm for 3D-HEVC system application,” Int. J. Performability Eng., vol. 14, no. 8, pp. 1712–1718, 2018.
[17] Y.-X. Song and K.-B. Jia, “Early merge mode decision for texture coding in 3D-HEVC,” J. Vis. Commun. Image Represent., vol. 33, pp. 60–68, Nov. 2015.
[18] J.-R. Lin, M.-J. Chen, C.-H. Yeh, Y.-C. Chen, L.-J. Kau, C.-Y. Chang, and M.-H. Lin, “Visual perception based algorithm for fast depth intra coding of 3D-HEVC,” IEEE Trans. Multimedia, early access, Mar. 31, 2021, doi: 10.1109/TMM.2021.3070106.
[19] Q. Zhang, H. Chang, X. Huang, L. Huang, R. Su, and Y. Gan, “Adaptive early termination mode decision for 3D-HEVC using inter-view and spatio-temporal correlations,” AEU - Int. J. Electron. Commun., vol. 70, no. 5, pp. 727–737, May 2016.
[20] Z. Pan, X. Yi, and L. Chen, “Motion and disparity vectors early determination for texture video in 3D-HEVC,” Multimedia Tools Appl., vol. 79, nos. 7–8, pp. 4297–4314, Feb. 2020.
[21] Q. Zhang, S. Wei, and R. Su, “Low-complexity texture video coding based on motion homogeneity for 3D-HEVC,” Sci. Program., vol. 2019, pp. 1–13, Jan. 2019.
[22] S. Bakkouri and A. Elyousfi, “FCM-based fast texture CU size decision algorithm for 3D-HEVC inter-coding,” in Proc. IEEE/ACS 17th Int. Conf. Comput. Syst. Appl. (AICCSA), Nov. 2020, pp. 1–5.
S. Bakkouri, A. Elyousfi, and H. Hamout, “Fast CU size and mode decision algorithm for 3D-HEVC intercoding,” Multimedia Tools Appl., vol. 79, pp. 6987–7004, Mar. 2020.

J. Chen, B. Wang, J. Liao, and C. Cai, “Fast 3D-HEVC inter mode decision algorithm based on the texture correlation of viewpoints,” Multimedia Tools Appl., vol. 78, no. 20, pp. 29291–29305, Oct. 2019.

Y. Li, G. Yang, Y. Zha, X. Ding, and R. Gong, “Probability model-based early merge mode decision for dependent views coding in 3D-HEVC,” ACM Trans. Multimedia Comput., Commun., Appl., vol. 14, no. 4, pp. 1–15, Nov. 2018.

P. Mukhopadhyay, An Introduction to the Theory of Probability, Singapore: World Scientific, 2012.

A. W. Marshall and I. Olkin, “A one-sided inequality of the Chebyshev type,” Ann. Math. Statist., vol. 31, no. 2, pp. 488–491, Jun. 1960.

Available: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16-0/.

K. Müller and A. Vetro, Common Test Conditions of 3D Core Experiments, document Rec. JCT3V-G1100, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, San Jose, CA, USA, Jan. 2014.

G. Bjontegaard, Calculation of Average PSNR Differences Between RD Curves, document ITU-T SG16/Q6 VCEG-M33, Austin, TX, USA, Apr. 2001.

G. Bjontegaard, Improvements of the BD-PSNR Model, document ITU-T SG16/Q6, VCEG-A111, Berlin, Germany, Jul. 2008.

MEI-JUAN CHEN (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Sun Yat-sen University, Kaohsiung, Taiwan, in 1997, 1999, and 2002, respectively. From 2000 to 2006, she was an Assistant Professor with the Department of Electrical Engineering, National Sun Yat-sen University. From 2006 to 2010, she held 47 patents in the USA, Taiwan, and China. His research interests include multimedia, video communication, three-dimensional reconstruction, video coding, image or video processing, and big data. He became a fellow of IET, in 2017. He is an active TC Member of the IEEE Communication Society on Multimedia Communication, APSIPA, and IWAIT. He is also one of the founding members of the ACM SIGMM Taiwan Chapter. He is on the Best Paper Award Committee of JVCI and APSIPA. He was a recipient of the 2007 Young Researcher Award of NSYSU, the 2011 Outstanding Young Electrical Engineer Award from the Chinese Institute of Electrical Engineering, the 2013 Distinguished Young Researcher Award of NSYSU, the 2013 IEEE MMSP Top 10% Paper Award, the 2014 IEEE GCCE Outstanding Poster Award, the 2015 APSIPA Distinguished Lecturer, the 2016 NAR Labs Technical Achievement Award, the Superior Achievement Award, the 2017 IEEE SPS Tainan Section Chair, the 2017 Distinguished Professor Award of NTNU, the IEEE Outstanding Technical Achievement Award (IEEE Tainan Section), and the Second Best Paper Award from 2020 International Computer Symposium. He was an Associate Editor of the EURASIP Journal on Advances in Signal Processing, and APSIPA Transactions on Signal and Information Processing.

JIE-RU LIN received the B.S., M.S., and Ph.D. degrees from the Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan, in 2015, 2016, and 2020 respectively. From August 2018 to January 2019 and from August 2019 to January 2020, he was a Lecturer with the Department of Electrical Engineering, National Dong Hwa University. He is currently an Engineer with the Industrial Technology Research Institute, Taiwan. His research interests include video coding and point cloud compression. He was a recipient of the Best Paper Award from the National Symposium on Telecommunication, in 2014, the Best Paper Award from IPPR Society and the Excellent Master Thesis Award from the Taiwan Institute of Electrical and Electronic Engineering, in 2016, the Best and Excellent Master Thesis Supervision Award from the Chinese Institute of Electrical Engineering, Taiwan, in 2020, and the Second Best Paper Award from 2020 International Computer Symposium.
MIN-HUI LIN received the B.S. degree in electrical engineering from National Kaohsiung University, Kaohsiung, Taiwan, in 2016, and the Ph.D. degree in electrical engineering from National Sun Yat-sen University, Kaohsiung, in 2021. She is currently a Senior Engineer with Qualcomm Inc., Taiwan. Her research interests include the deep learning for computer vision, 3D reconstruction, and multimedia applications. She was a recipient of the Best Oral Presentation Award from the 2018 IEEE Asia-Pacific Conference on Circuits and Systems and the Second Best Paper Award from the 2020 International Computer Symposium.

LIH-JEN KAU (Senior Member, IEEE) received the B.S. degree in control engineering, and the M.S. and Ph.D. degrees in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1991, 1997, and 2008, respectively. From 1996 to 1998, he was with Chunghwa Telecom, Taipei, Taiwan, where he was a Senior Technician. In August 1998, he joined the Department of Computer and Communication Engineering, Dahan Institute of Technology, Hualien, Taiwan, as an Instructor, where he has been an Assistant Professor, since February 2006. Since August 2009, he has been with the Department of Electronic Engineering (EE), National Taipei University of Technology (NTUT), Taipei. He is currently an Associate Professor with the EE Department, NTUT. His research interests include image or video coding and processing, healthcare and biomedical information technology, and autonomous vehicle system integration. He was a recipient of the 2003 ZyXEL Scholarship Award from ZyXEL Communications Corporation, Hsinchu, the Best Live Demonstration Award from the 2016 IEEE Biomedical Circuits and Systems Conference, and the Best Paper Award from the 2019 IEEE International Conference of Intelligent Applied Systems on Engineering.

CHUAN-YU CHANG (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2000. From 2009 to 2011, he was the Chair of the Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology (YunTech), Taiwan. From 2011 to 2019, he was the Dean of the Research and Development and the Director of the Incubation Center for Academia-Industry Collaboration and Intellectual Property, YunTech. He is currently the Deputy General Director of the Service Systems Technology Center, Industrial Technology Research Institute, Taiwan. He is also a Distinguished Professor with the Department of Computer Science and Information Engineering, YunTech. He has authored or coauthored more than 200 publications in journals and conference proceedings in his research fields, which include computational intelligence and their applications to medical image processing, automated optical inspection, emotion recognition, and pattern recognition. He is a Life Member of IPPR and TAAI and an IET Fellow. He was the Program Co-Chair of TAAI 2007, CVGIP 2009, 2010–2019 International Workshop on Intelligent Sensors and Smart Environments, and the third International Conference on Robot, Vision and Signal Processing (RVSP 2015). He was the General Co-Chair of the 2012 International Conference on Information Security and Intelligent Control, 2011–2013 Workshop on Digital Life Technologies, CVGIP 2017, WIC 2018, ICS 2018, and WIC 2019. From 2015 to 2017, he was the Chair of the IEEE Signal Processing Tainan Chapter and the Representative for Region Ten of IEEE SPS Chapters Committee. He is also the President of the Chinese Image Processing and Pattern Recognition Society and the Taiwan Association for Web Intelligence Consortium.