IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 6, 2019, accepted January 28, 2019, date of publication February 4, 2019, date of current version March 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897161

Bi-Layer Texture Discriminant Fast Depth Intra Coding for 3D-HEVC

JIABAO ZUO[1,2], JING CHEN[1,2], (Member, IEEE), HUANQIANG ZENG[1,2], (Senior Member, IEEE), CANHUI CAI[1,2], (Senior Member, IEEE), AND KAI-KUANG MA[1,3], (Fellow,+ IEEE)

[1]School of Information Science and Engineering, Huaqiao University, Xiamen 361021, China
[2]Xiamen Key Laboratory of Mobile Multimedia Communications, Xiamen 361021, China
[3]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

Corresponding author: Jing Chen (chenjing8005@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61802136 and Grant 61871434, in part by the Natural Science Foundation for Outstanding Young Scholars of Fujian Province under Grant 2019J06017, in part by the Natural Science Foundation of Fujian Province under Grant 2017J05103 and Grant 2016J01308, in part by the Fujian-100 Talented People Program, Teaching Research Project of Fujian Province (FBJG20180038), in part by the High-level Talent Innovation Program of Quanzhou City under Grant 2017G027, in part by the Promotion Program for Young and Middle-Aged Teacher in Science and Technology Research of Huaqiao University under Grant ZQN-YX403, and in part by the High-Level Talent Project Foundation of Huaqiao University Grant 16BS709, Grant 14BS201, and Grant 14BS204.

ABSTRACT 3D-HEVC explores several new intra-frame depth coding tools to improve the coding efficiency, but at the price of high computational complexity, which hinders the practical applications of 3D-HEVC. Therefore, a bi-layer texture discriminant fast depth intra coding for 3D-HEVC is proposed in this paper. With a sum of gradient matrix (SGM), the texture complexity (TC) of current depth coding unit (CU) and its subblocks are calculated at once. The values of TCs indicate whether to split the depth CU further and skip unnecessary DMMs checking. The experimental results show that compared with the original 3D-HEVC, the proposed algorithm reduces the depth encoding time by 44.8%, and the average bit rate of the synthetic viewpoint is only increased by 0.38%, which outperforms the state-of-the-art fast 3D-HEVC algorithms.

INDEX TERMS Depth intra coding, 3D-HEVC, CU size decision, fast mode decision.

I. INTRODUCTION

For 3D video services, such as 3D television, 3D movies, 3D games, and so on, the immerse feeling of the scene to the users is in highly demanding. Consequently, the representation and the coding scheme of the immerse media ''3D video'' have achieved more and more attentions. For that, some formats for 3D video are widely-used, including conventional stereo video (CSV), multi-view video (MVV), and multi-view video plus depth (MVD) [1]–[3]. Among them, the MVD is composed of two sequences for each view, one is the 2D color sequence and the other is its corresponding depth map sequence. Compared with MVV, fewer viewpoints needed to be captured and coded; Compared with CSV, depth information of the scene represented by MVD is more accurate. For 3D video coding, 3D extension standard of High Efficiency Video Coding (3D-HEVC) has been established by Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) [4], which represents 3D video by MVD format. In 3D-HEVC, the color and depth sequences of MVD are encoded by 3D video encoder, the resulting bitstream can be adaptively decoded by 3D video decoder, stereo video decoder and traditional 2D video decoder for the sake of need [5]–[6]. The depth-image-based-rendering (DIBR) technique enables to synthesize more intermediate virtual views that are appropriate for displaying [7].

In 3D-HEVC, the coding processes for color and depth sequences are similar, which inhere the quad-tree coding structure of HEVC. Although there are several tools for 3D-HEVC to achieve better coding performance, the complexity is relatively high, especially for depth map sequence coding. The depth map indicates the distance of objects from the camera, which consists a large number of flat areas and few edges. Compared with flat area, the edge information is more essential and needs to be coded more precisely [8]. 3D-HEVC introduces several new prediction modes and

The associate editor coordinating the review of this manuscript and approving it for publication was Mohsin Jamil.

coding tools to preserve sharp object edges in depth map, such as depth modeling modes (DMM) [9], segment-wise depth coding (SDC) [10], and motion parameter Inheritance (MPI) [11]. These new modes and tools with the quad-tree coding structure result in high computational load, which limits the practical application of 3D-HEVC. Therefore, the study on fast algorithm for 3D-HEVC to reduce the complexity without sacrificing the rate-distortion (RD) performance is necessary and valuable.

There are some researches that focus on fast intra coding for depth map and achieve remarkable results. In [12], two fast intra mode decision strategies are proposed by utilizing the depth map characteristics to predict the depth intra mode of current coding unit (CU) and skip some unnecessary modes checking. In [13], evaluation results for intra conventional modes (ICMs) are utilized to determine whether DMMs can be skipped or not, and simplify the DMMs searching. By studying the correlation between the corresponding texture video CU and the current depth map CU, an optimal coding level parameter of depth map CU is defined in [14], which adaptively omits the unnecessary coding levels searching and computation in 3D-HEVC. In [15], the smoothness of current coding block can be estimated by the definition of sum-of-gradient, which enables to skip unnecessary DMMs checking for flat areas. Moreover, the partitioning size of CU can be prejudged by the sum-of-gradient, which saves the iterative partitioning and searching process. Reference [16] proposes a low complexity intra mode selection algorithm. By classifying intra modes into three activity classes for non-directional modes, directional modes and DMMs, respectively, each class is assigned with different mode-weight factors (0, 1, and 2). An intra mode complexity (IMC) parameter of current CU is defined to assign corresponding candidate intra modes for diverse CUs. The adaptive assignment of the candidate intra prediction modes speeds up the best mode decision process. And the early termination of intra prediction on unnecessary sizes saves the computational load a lot. By analyzing the correlation between texture and depth map, the work in [17] shows that when the corresponding texture CU is encoded by Skip mode, the current depth map CU has a high probability (over 95%) of being encoded by Skip or DIS. Then, a block-level fast algorithm is proposed, when Skip and DIS modes have a high chance of being selected, avoiding the remaining best mode searching procedure. Reference [18] propose a model-level scheme for reducing the most time-consuming steps at the intra prediction, with DMM-1 fast pattern selector (DFPS), pattern and gradient-based mode one Filter (P&GMOF) and enhanced depth rough mode decision (ED-RMD) algorithms, the time consuming of depth map intra prediction is saved. In summary, these fast algorithms can be summarized into two sorts, one is to skip some unnecessary modes checking, the other one is to terminate the CU partitioning process earlier.

Though there are lots of time saving achieved, the coding complexity of 3D-HEVC is relatively high. For that, a fast algorithm for depth intra coding based on bi-layer texture analysis is proposed in this paper, which predicts bi-layer texture at one time with low coding complexity. First, a texture complexity (TC) parameter is developed, and then, with TCs of current CU and its subblocks, the proposed method terminates the CU size splitting earlier and skips checking DMMs, which achieves salient time saving with almost the same coding efficiency. Since the TCs of current CU and corresponding subblocks are calculated at one time, the discriminant is with low complexity. Experimental results have shown that the proposed method is able to significantly reduce the computational complexity of depth intra coding in 3D-HEVC.

The remaining part of this paper is as follows. Section II shows the analysis of depth intra coding in 3D-HEVC. Section III illustrates the bi-layer texture discriminant strategy. The bi-layer texture discriminant fast depth intra coding algorithm is proposed in Section IV. And the experimental results and conclusions are presented in Sections V and VI, respectively.

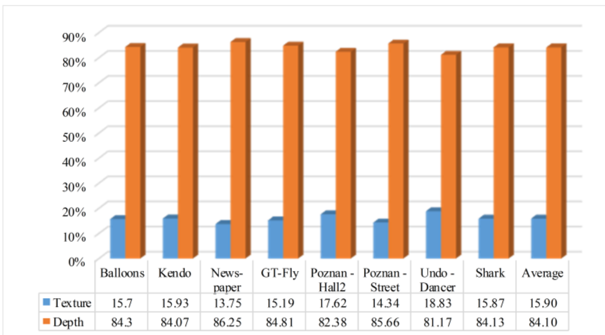

FIGURE 1. Encoding time occupation in 3D-HEVC.

II. THE ANALYSIS OF DEPTH INTRA CODING IN 3D-HEVC

A. THE ENCODING TIME OCCUPATION ANALYSIS

As an extension of HEVC, 3D-HEVC inherit the quad-tree coding structure for either texture or depth map sequence. Since there are a large number of flat areas in depth map, the edge information tends to be more important and needs to be coded precisely. Though introducing several new prediction modes and coding tools to preserve edges in depth map, the depth map coding is relatively time consuming. To analyze the percentage of encoding time cost, an experiment is conducted by using HTM 16.0 [19], with all intra configurations, and CU depth level ranges from 0 to 3. The encoding time occupations of eight sequences (Balloons, Kendo, Newspaper, GT-Fly, Poznan-Hall2, Poznan-Street, Undo-Dancer, Shark) are shown in Figure 1. It can be seen that, the coding time of the depth map sequence is more than four times of the corresponding texture one, which occupies over 84% of the total encoding time in average. Therefore, how to reduce the encoding time of the depth map without affecting the quality is vital for the practical application of 3D video services.

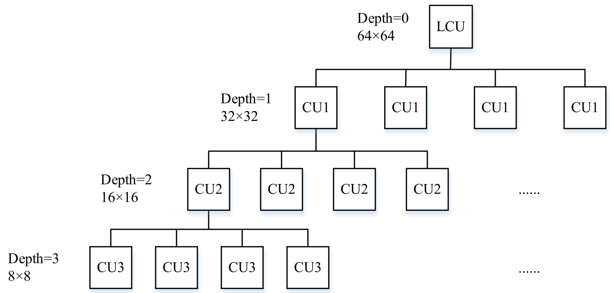

FIGURE 2. **Quad-tree splitting structure of CTU.**

B. THE TIME-CONSUMING DEPTH MAP CODING

1) QUAD-TREE SEGMENTATION PROCESS OF DEPTH CU

As shown in Figure 2, for each basic coding tree unit (CTU), the largest size of coding unit (CU) is 64×64, named LCU. It can be split into four smaller CUs in one segmentation depth. With four levels of segmentation depth (0-3), CU can be recursively split until the minimum CU size of 8×8. Therefore, there are four CU depth levels (CUDL), namely, 0, 1, 2, and 3, with size of 64×64, 32×32, 16×16 and 8×8, respectively. For each CTU, the optimal CUDL is selected by the rate distortion cost (RdCost) of each CU, which means the RdCost of each CU must be calculated first, and then, the CUDL with the smallest RdCost will be chosen as the optimal one. With the quad-tree segmentation process of CU, the recursive best CUDL searching process is obviously time consuming. If there is a way to predict the CUDL in advance, the complexity of the depth map coding can be effectively reduced.

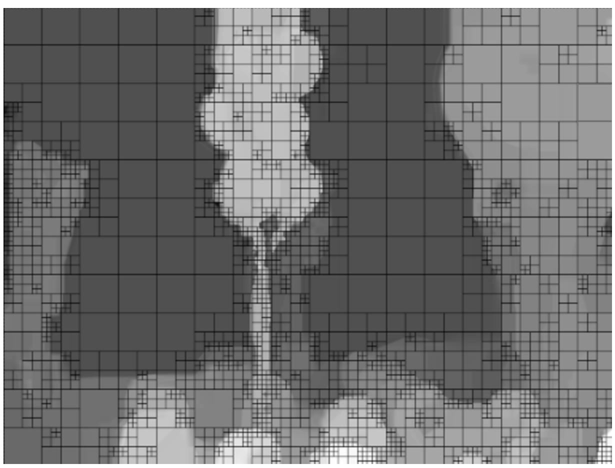

FIGURE 3. **The CU splitting result of depth map in "Balloons".**

As shown in Figure 3, the CU segmentation result of the first frame in depth intra coding of sequence "Balloons" is with larger size for flat area and smaller size for edge region. It indicates that the optimal CUDL has highly correlation with the texture complexity of CU. Furthermore, to study the distribution of optimal CUDL, four sequences, Balloons, Kendo, Poznan-Hall2 and Poznan-Street, are tested. As can be seen from Figure 4, there is a large percentage of selecting smaller CU segmentation depth. If the CUDL can be predicted by the characteristics of CU in an accurate manner, the quad-tree segmentation process and the iterative searching can be terminated earlier.

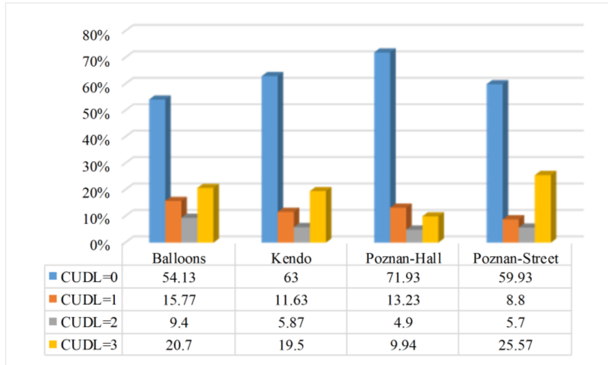

FIGURE 4. **CU depth level (CUDL) distribution for depth maps (%).**

TABLE 1. **Intra prediction modes for depth PU.**

Mode number	Mode
0	Intra-Planar
1	Intra-DC
2-34	Intra-Angular
37-38	Intra-DMMs

2) THE OPTIMAL INTRA MODE SELECTION FOR DEPTH PU

In order to achieve high coding efficiency, 3D-HEVC provides 37 intra prediction modes for depth PU, including 35 traditional HEVC modes and two DMM modes, as listed in Table I. The selection process of the optimal intra prediction mode is as follows. (1) the rough mode decision (RMD) algorithm is employed to select several HEVC intra mode candidates roughly by minimizing the rate distortion (RD) cost, calculate by formula (1).

$$R_{\cos t} = D_{Had} + \lambda \times R_{mode} \tag{1}$$

where *DHad* represents the absolute sum of Hadamard transformed residual for a block, λ is the Lagrange coefficient, and *Rmode* is the number of bits for current prediction mode [16]; (2) Most probable mode (MPM) and DMM modes are added to the candidate list; (3) Find out the optimal mode from the obtained candidate list with the minimum cost of mode decision, which is specified by formula (2).

$$J_{mode} = (SSE_{luma} + \omega_{chroma} \cdot SSE_{chroma}) + \lambda \times B_{mode} \tag{2}$$

where *SSE* represents the sum of squared error between the current CU and the matching block, *ωchroma* is the weighting factor for chroma components, and *Bmode* is the bitrate cost dependents on each diverse decision case [16].

Since there are a large number of flat areas in depth map, the percentage of selecting DMMs is relatively small, but

$I(x\text{-}1, y\text{-}1)$	$I(x, y\text{-}1)$	$I(x\text{+}1, y\text{-}1)$
$I(x\text{-}1, y)$	$I(x, y)$	$I(x\text{+}1, y)$
$I(x\text{-}1, y\text{+}1)$	$I(x, y\text{+}1)$	$I(x\text{+}1, y\text{+}1)$

FIGURE 5. The definition of texture unit.

the DMM mode checking process is time consuming. If the DMM mode checking can be skipped for those flat areas, the complexity of the algorithm can be reduced, with no harm to the coding efficiency.

III. PROPOSED BI-LAYER TEXTURE DISCRIMINANT

A. THE SUM OF GRADIENT MATRIX (SGM)

To evaluate the texture complexity of the depth map, a sum of gradient matrix (SGM) is proposed. First, a texture unit [20] is defined as Figure 5 shows, the unit size is of 3×3, where $I(x, y)$ denotes the current pixel point. The gradient of $I(x, y)$ in four directions (0°, 90°, 45°, and 135°) are calculated by equations (3)-(6), respectively.

$$A_1(x, y) = |I(x-1, y) - I(x+1, y)| \tag{3}$$
$$A_2(x, y) = |I(x, y-1) - I(x, y+1)| \tag{4}$$
$$A_3(x, y) = |I(x+1, y-1) - I(x-1, y+1)| \tag{5}$$
$$A_4(x, y) = |I(x-1, y-1) - I(x+1, y+1)| \tag{6}$$

Then, for each CU, the sum of gradient matrix is generated. To make it simple, the gradients of CU edge pixels do not need to be calculated. Let CU be N×N, the generated SGM is of size W×W, where W equals to N-2. The value of each matrix element is obtained by $P(m, n)$, which denotes the sum of four direction gradients for pixels in CU as shown in equation (7), where, (m, n) denotes the position of the texture unit in CU, and $m \in (1, N-2)$, $n \in (1, N-2)$.

$$P(m, n) = \sum\nolimits_{i=1}^{4} Ai(m, n) \tag{7}$$

The illustration of depth CU and the corresponding SGM is shown in Figure 6. Figure 6 (a) is an example of flat CU and its SGM. Figure 6 (b) is a CU with complex texture and its SGM. It can be seen that the element values of SGM reflect the texture detail well.

B. TEXTURE COMPLEXITY OF CURRENT CU AND ITS SUBBLOCKS

As illustrated in Figure 3, the optimal splitting depth of current CU is related to the texture complexity. Specifically, the areas with complex textures are usually split into smaller-sized CUs, while the region with smooth texture is always coded with larger-sized CUs. Therefore, the current CU can be classified into Split-CU (SCU) and Non-Split-CU

80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80

→

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

(a)

19	19	19	20	26	35	45	55
19	19	19	20	26	35	44	54
19	19	19	19	26	35	43	52
19	19	19	19	26	35	43	52
21	21	19	17	29	37	44	52
31	38	37	35	40	44	48	52
32	38	36	32	41	45	48	53
40	52	54	52	52	53	53	54

→

0	2	22	46	54	58
0	1	22	47	52	55
4	4	26	55	51	50
51	57	48	59	49	45
51	52	47	54	38	35
44	54	51	51	33	24

(b)

FIGURE 6. The depth CU and the corresponding SGM. (a) Flat CU and its SGM. (b) Complex texture CU and its SGM.

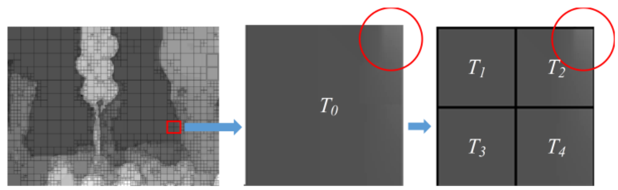

FIGURE 7. The TCs of particular CU and it's subblocks ($T_0 = 5830$, $T_1 = 188$, $T_2 = 4104$, $T_3 = 186$ and $T_4 = 1352$).

(NCU). If the current CU belongs to SCU, it needs to be split into smaller CUs; if the current CU is an NCU, the splitting process of CU is terminated.

For current CU, the texture complexity (TC) can be denoted as T_0, as follows:

$$T_0 = \sum_{n=1}^{N-2} \sum_{m=1}^{N-2} P(m, n) \tag{8}$$

Note that the CU with small T_0 may have complex texture in its subblocks, as illustrated in Figure 7, subblock with rich textures is circled in red.

In this case, if the CU size is determined without checking the texture complexity of its subblocks, the CU size prediction may be inaccurate, leading to the reduction of the coding efficiency. Therefore, The texture complexity of four subblocks of current CU are also examined by T_1, T_2, T_3 and T_4, which calculated by the same way with T_0.

IV. BI-LAYER TEXTURE DISCRIMINANT FAST DEPTH INTRA CODING FOR 3D-HEVC

A. FAST CU SIZE DECISION

1) EARLY DETERMINATION OF NCU

With T_0, T_1, T_2, T_3 and T_4, the texture complexities (TCs) of current CU and its subblocks are determined. If the current CU can be predicted as an NCU, the CU splitting process can be terminated to save the coding time. Therefore, an early determination of NCU with bi-layer texture discriminant is proposed. The range of TC is divided into two discriminants:

$$\text{Discriminant 1: } 0 \leq T_0 < Th_{Ni}$$
$$\text{Discriminant 2: } \begin{cases} Th_{Ni} < T_0 < ave_i \\ T_1 < 0.5 \times T_0 \\ T_2 < 0.5 \times T_0 \\ T_3 < 0.5 \times T_0 \\ T_4 < 0.5 \times T_0 \end{cases} \quad (9)$$

where i denotes the splitting depth of current CU, Th_{Ni} is the threshold for NCU, ave_i represents the TC average value of corresponding CU in the training frame. For discriminant 1, the range of T_0 is from 0 to Th_{Ni}, which means the TC of current CU is small and tends to be a flat area. Then, the CU is directly classified into NCU, and the splitting process can be stopped earlier. For discriminant 2, when T_0 is in the range from Th_{Ni} to ave_i, the TC of subblocks should be examined, in case that texture concentrates on the subblocks. Thus, besides T_0 belongs to the range from Th_{Ni} to ave_i, when the TCs of corresponding subblocks T_1, T_2, T_3 and T_4 are all less than 1/2 of T_0, the current CU can be considered as NCU and do not need to split further.

Determination of the NCU Threshold:

The value of Th_{Ni} is related to the training frame. The initial training frame is set to be the first frame of the depth map. The threshold value is calculated with formula (10).

$$Th_{Ni} = \alpha \times ave_i \quad (10)$$

where α is the mean coefficient range from 0 to 1, and CU depth i is 0, 1, and 2. Considering the temporal correlation of the coding video, the training frame is refreshed every 30 frames.

Since the value of Th_{Ni} is associated with ave_i and α, how to choose a reasonable value of α is essential. With experiment tested in four sequences (GT-Fly. Poznan-Street, Shark, Undo-Dancer), by varying α from 0.6 to 0.8 and 1.0, the hit-rate of the CU terminate segmentation is shown in Figure 8. The calculation of the hit-rate is as formula (11).

$$hit-rate = \frac{Num_N CU(0 \leq T_0 < Th_{Ni})}{Num_S CU(0 \leq T_0 < Th_{Ni}) + Num_N CU(0 \leq T_0 < Th_{Ni})} \quad (11)$$

where $Num_N CU(0 \leq T_0 < Th_{Ni})$ indicates the number of CUs that stop segmentation when the TC of current CU is smaller than the threshold, while $Num_S CU(0 \leq T_0 < Th_{Ni})$ indicates the number of CUs which still split when the TC of current CU is smaller than the threshold.

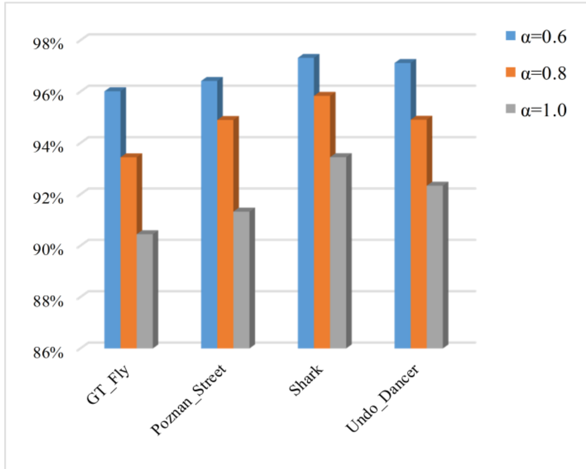

FIGURE 8. The hit-rate of diverse α.

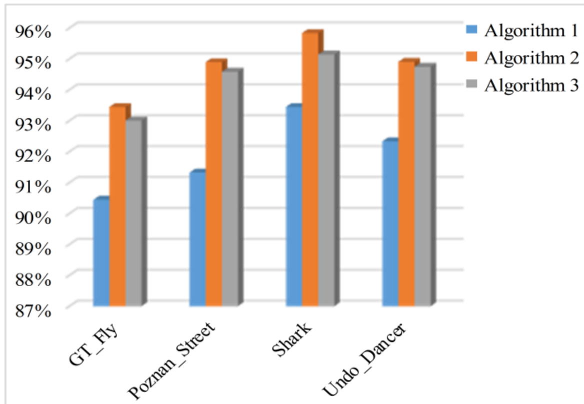

FIGURE 9. The hit-rate of three algorithms.

As can be seen from Figure 8 that the threshold increases with α, and the hit-rate is significantly reduced. Considering the tradeoff between distortion and the encoding time, $\alpha = 0.8$ is empirically-determined.

To verify the effectiveness of the proposed bi-layer texture discriminant fast depth intra coding algorithm, the comparisons with one-layer texture discriminant are shown in Figure 9. There are three fast depth intra coding algorithms as follows.

Algorithm 1: $\alpha = 1.0$, one-layer texture discriminant, the range of TC: $0 \leq TC < ave_i$;

Algorithm 2: $\alpha = 0.8$, one-layer texture discriminant, the range of TC: $0 \leq TC < 0.8 \times ave_i$;

Algorithm 3: $\alpha = 0.8$, bi-layer texture discriminant, the range of TC: $0 \leq TC < ave_i$.

It can be seen from Figure 9 that the hit-rate of Algorithm 2 is higher than that of Algorithm 1, but this is at the expense of reducing the threshold range. The threshold range of Algorithm 3 equals to that of Algorithm 1, and the hit rate is similar to Algorithm 2, which indicates that Algorithm 3 can balance the efficiency and the correctness of TC prejudgment.

In addition, the comparison of two algorithms (Algorithm 2 and Algorithm 3) is shown in Table II. Where, "Synth

PSNR/Total bitrate" denotes the BD-rate results for synthesized views. "Time Saving" denotes the time saving compared with the original 3D-HEVC. It can be noticed that Algorithm 3 can achieve more time saving compared with Algorithm 2, up to about 25% in "GT-Fly". The average incensement of time saving is about 9%. And the average bit rate of the synthetic viewpoint is only increased by 0.03%, which demonstrates the effectiveness of the proposed bi-layer texture discriminant fast algorithm.

2) EARLY DETERMINATION OF SCU

As analysis, with the TCs of current CU and its subblocks, the CU segmentation process can be terminated earlier. On the contrary, if the current CU is full of details, the current coding process can be skipped, and turn to the next CU depth directly. Thus, an early determination of SCU with bi-layer texture discriminant is proposed. The range of TC is divided into two segments.

$$\begin{aligned} &Discriminant1: \ T_0 > Th_{S2i} \\ &Discriminant2: \begin{cases} Th_{S1i} < T_0 < Th_{S2i} \\ (T_1 or T_2 or T_3\ T_4) > 0.5 \times T_0 \end{cases} \end{aligned} \quad (12)$$

where $ThS1i$ and $ThS2i$ are two thresholds for CU depth equals to i, and $ThS2i$ is larger than $ThS1i$. For discriminant 1, the range of $T0$ is larger than $ThS2i$, which means the high complexity of current CU. Then, the current CU is directly classified into SCU, then, skip the prediction process of current depth and split into smaller CUs. For discriminant 2, when $T0$ is in the range from $ThS1i$ to $ThS2i$, the TC of subblocks should be examined, in case that texture concentrates on the subblocks. Thus, besides $T0$ varies from $ThS1i$ to $ThS2i$, when one of TCs of corresponding subblocks (T_1, T_2, T_3, T_4) is more than 1/2 of $T0$, the current CU can be considered as an SCU, which can be split further directly.

Determination of the SCU threshold:

The value of $ThSni$ is set according to the training frame, calculated by (13). Where, n equals to 1 or 2, which indicates two thresholds, respectively.

$$Th_{Sni} = \beta \times avei \quad (13)$$

The value of $ThSni$ is associated with β and ave_i. Where, ave_i is the TC average value of corresponding CU in the training frame. A reasonable β is important. With experiment tested in GT-Fly, Poznan-Street, Shark, Undo-Dancer sequences, by varying β from 1.0 to 2.0, the hit-rate of the CU terminate segmentation is shown in Figure 10.

It can be seen that the threshold increases with β, and the hit-rate is significantly increased. When $\beta = 2.0$, the correctness hit-rate of CU segmentation can achieve 95% or more, which indicates that if the $T0$ of current CU is larger than the threshold at this time, the accuracy of classifying CU into SCU is very high, and the distortion will be small. Therefore, the threshold $ThS2i$ is set in the case of $\beta = 2.0$. When $\beta = 1.5$, the accuracy of hit-rate is between

TABLE 2. Performance comparison for one-layer texture and bi-layer texture discriminant fast algorithm with equal α (%).

Sequences	Algorithm 2		Algorithm 3	
	Synth PSNR /Total bitrate	Time Saving	Synth PSNR /Total bitrate	Time Saving
GT-Fly	0.04	25.73	0.12	50.67
Poznan-Street	0.02	31.15	0.08	33.99
Shark	0.04	37.29	0.04	37.91
Undo-Dancer	0.08	41.19	0.07	45.54
Average	**0.05**	**33.84**	**0.08**	**42.03**

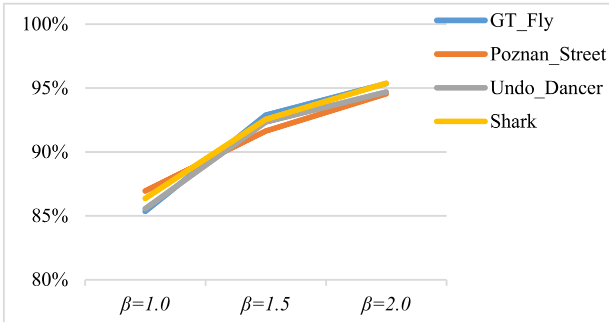

FIGURE 10. The hit-rate of diverse β.

91% to 93%, and when β equals to 1.0, the hit-rate declines significantly. Then, the threshold $ThS1i$ is set when $\beta = 1.5$.

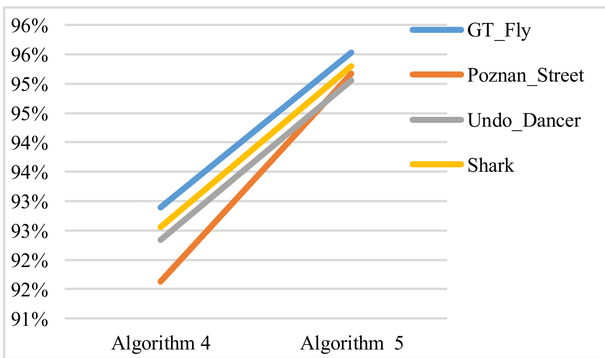

FIGURE 11. The hit-rate of two algorithms ($\beta = 1.5$).

In order to increase the hit-rate of the segmentation when $\beta = 1.5$, the TC of subblocks should also be examined. To substantiate the effectiveness of the proposed bi-layer texture discriminant for SCU, the comparison with one-layer texture discriminant is illustrated in Figure 11. There are two fast depth intra coding algorithms as follows.

Algorithm 4: $\beta = 1.5$, one-layer texture discriminant, the range of TC: $TC \geq 1.5 \times ave_i$;

Algorithm 5: $\beta = 1.5$, bi-layer texture discriminant, the range of TC: $TC \geq 1.5 \times ave_i$.

It can be seen from Figure 11 that the hit-rate of Algorithm 5 is higher than that of Algorithm 4, which indicates that the bi-layer texture discriminant is superior to the one-layer texture ones.

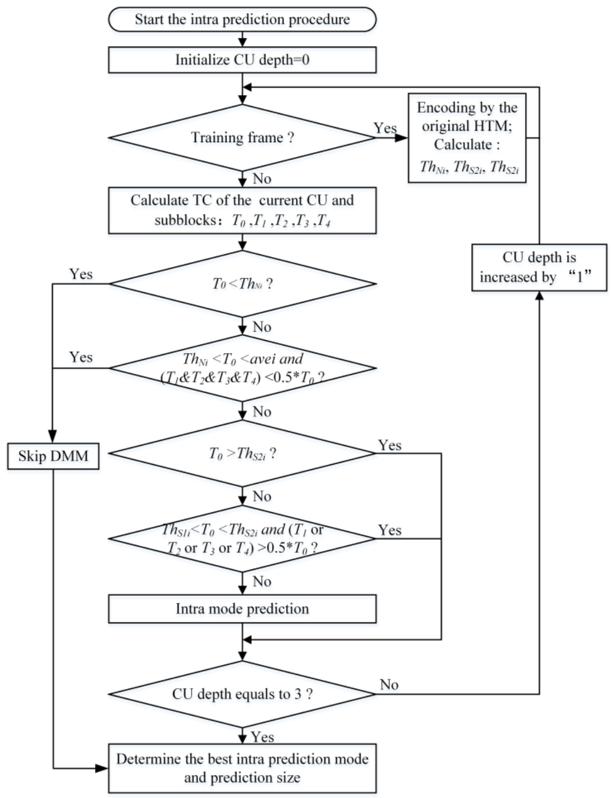

FIGURE 12. Flowchart of the proposed algorithm for each CTU.

TABLE 3. Test configurations.

Input video sequence:	3 viewpoints
All-intra configuration	
CU Depth:	0-3
DMM:	ON
Texture QP values:	25, 30, 35, 40
Depth QP values:	34, 39, 42, 45
VSO:	ON
Texture SAO: ON	Depth SAO: OFF
RDOQ:	ON

B. DMMs SKIPPING

Considering that the optimal intra prediction mode of CU has a strong correlation with the texture complexity, with bi-layer texture discriminant, the DMM modes, which particularly designed for edge regions, are in a small probability be chosen as the optimal intra prediction mode for those flat CUs. Thus, the time-consuming DMMs checking process can be skipped. The specific decision is as follows: if TCs of current CU and its subblocks are within the threshold range set by formula (9), then the checking of DMMs is skipped to further save the computational complexity.

C. FLOWCHART OF THE PROPOSED ALGORITHM

According to the above analysis, a bi-layer texture discriminant fast depth intra coding algorithm is proposed. The flowchart is illustrated in Figure 12. For each CTU, the intra prediction procedure of the current CU is as follows.

(1) Initialize CU depth=0;

(2) Determine whether the current frame is a training frame. If yes, the training frame is encoded by the original HTM, and each threshold required is calculated $ThNi$, $ThS1i$ and $ThS2i$. Else, turn to (3);

(3) Calculate the TC of current CU $T0$ and TCs of its subblocks ($T_1 \sim T_4$).

(4) Early determination of NCU: ① Compare $T0$ and $ThNi$, if $T0$ is less than $ThNi$, the CU splitting process is terminated; ② If $T0$ is between $ThNi$ and $avei$, TCs of its subblocks are all less than half of the TC of current CU ($T_1 \sim T_4 < 0.5 \times T0$), stop the segmentation of current CU. For these cases, skip DMMs checking, and turn to (7); Else, turn to (5);

(5) Early determination of SCU: ① If the TC of current CU ($T0$) is larger than $ThS2i$, the current CU is classified as SCU; ② If the TC of current CU is between $ThS1i$ and $ThS2i$ ($ThS1i < T0 < ThS2i$), and any TC of subblocks TC is larger than half of the TC of current CU, then, the current CU is in a region with details. For these SCUs, skipping the intra prediction process of current CU, turn to (6); Else, to do the intra mode prediction step by step, after that, turn to (6);

(6) If the current CU depth equals to 3, then, turn to (7); Else, CU depth+1, turn to (2);

(7) Determine the optimal intra prediction mode and CU size.

TABLE 4. Test sequences and configuration parameters.

Resolution	Test sequences	Total frames
1024×768	Balloons	300
1024×768	Kendo	300
1024×768	Newspaper	300
1920×1088	GT-Fly	200
1920×1088	Poznan-Hall2	250
1920×1088	Poznan-Street	250
1920×1088	Undo-Dancer	250
1920×1088	Shark	300

TABLE 5. Comparison between the proposed algorithm and 3D-HEVC (%).

Test sequence	V/V	V/T	S/T	ETS_{total}	ETS_{depth}
Balloons	0.00	0.07	0.31	31.2	37.0
Kendo	0.00	0.07	-0.10	33.9	40.8
Newspaper	0.00	-0.03	0.61	31.3	36.2
GT-Fly	0.00	0.03	0.19	44.5	53.1
Poznan-Hall2	0.00	-0.05	1.30	42.6	54.4
Poznan-Street	0.00	-0.09	0.27	31.6	37.1
Undo-Dancer	0.00	0.00	0.25	41.1	53.7
Shark	0.00	0.06	0.19	38.1	46.1
1024×768	**0.00**	0.04	0.27	**32.1**	**38.0**
1920×1088	**0.00**	**-0.01**	**0.44**	**39.6**	**48.9**
Average	**0.00**	**0.01**	**0.38**	**36.8**	**44.8**

V. EXPERRMENTAL RESULTS

To evaluate the effectiveness of the proposed algorithm, experiments are performed in 3D-HEVC reference software,

TABLE 6. Comparison between the proposed and the state-of-the-art fast depth intra coding algorithms (%).

Test sequence	[17]		[18]		The proposed	
	S/T	ETS_{depth}	S/T	ETS_{depth}	S/T	ETS_{depth}
Balloons	0.21	26.5	0.36	34.1	0.31	37.0
Kendo	0.30	23.2	0.37	33.9	-0.10	40.8
Newspaper	0.26	23.2	0.46	35.4	0.61	36.2
GT-Fly	Training sequences		0.12	40.6	0.19	53.1
Poznan-Hall2	0.39	30.2	0.43	38.8	1.30	54.4
Poznan-Street	0.15	30.0	0.22	41.7	0.27	37.1
Undo-Dancer	Training sequences		0.12	38.5	0.25	53.7
Shark	0.13	28.4	0.11	36.2	0.19	46.1
1024×768	**0.26**	**24.3**	**0.40**	**34.5**	**0.27**	**38.0**
1920×1088	**0.22**	**29.5**	**0.2**	**39.2**	**0.44**	**48.9**
Average	**0.24**	**26.9**	**0.27**	**37.4**	**0.38**	**44.8**

HTM 16.0 [19], under the all-intra configuration cases suggested by the common test condition (CTC) [21] drafted by JCT-3V. The configuration parameters of the encoder are as shown in Table III. Each sequence has been encoded by four pairs of quantization parameters (T_QP, D_QP): (25, 34), (30, 39), (35, 42), (40, 45). Test sequences and parameters are listed in Table IV.

In our experiments, V/V (Video PSNR/Video bitrate) indicates the BD-rate coding result for the coded color sequences, only bitrate for color sequence is considered; V/T (Video PSNR/Total bitrate) denotes the BD-rate coding result for coded color and depth sequences, which reflects the quality of coded video; S/T (Synth PSNR/Total bitrate) denotes the BD-rate result for synthesized views with total bitrate of both color and depth sequences, which reflects the quality of synthesized views; *ETStotal* denotes the total time savings for encoding; *ETSdepth* denotes the time savings for depth map coding. Compared to the original HTM 16.0 encoder, *ETStotal* and *ETSdepth* are calculated by formula (14) [16].

$$ETS = \frac{Time_{HTM} - Time_{method}}{Time_{HTM}} \times 100\% \tag{14}$$

where $Time_{HTM}$ and $Time_{method}$ denote the coding time of the original HTM 16.0 encoder and the proposed algorithm, respectively.

A. COMPARISON WITH 3D-HEVC STANDARD

The coding performance and the time saving of the proposed algorithm with the comparison to the original 3D-HEVC (i.e., HTM 16.0), are shown in Table V.

It can be seen that the proposed algorithm can achieve 44.8% of depth coding time saving in average and 36.8% of total time saving with slightly quality degradation of synthesized sequences, which demonstrates the effectiveness of the proposed algorithm. It can also be noticed that the proposed algorithm is more suitable for sequences with large global or local motions, like ''GT-Fly''. For those sequence with large smooth area or with small/medium local motion, like ''Poznan-Hall2'', the proposed may cause distortions.

B. COMPARISON WITH THE STATE-OF-ART FAST 3D-HEVC ALGORITHMS

The coding efficiency and the time-consuming performance of the proposed algorithm compared with the state-of-the-art fast depth intra coding algorithms for 3D-HEVC [17] and [18] are listed in Table VI. It can be seen that the proposed can achieve more 17.9% and 7.4% of time-saving, respectively, with almost the same average BD-rate results, which indicates the outperformance of the proposed fast algorithm foe 3D-HEVC.

VI. CONCLUSION

To reduce the coding complexity of 3D-HEVC depth map, a bi-layer texture discriminant fast depth intra coding algorithm for 3D-HEVC is proposed. With bi-layer texture discriminant, current CU can be early classified into Split CU (SCU) and Non-split CU (NCU). If CU belongs to an NCU, the CU splitting is early terminated and DMMs checking is also early skipped; If CU belongs to an SCU, the proposed method will skip the prediction mode searching for current CU, and turn to the next CU splitting depth. Experimental results show that the proposed algorithm can achieve 44.8% of coding time saving on average while keeping almost the same coding efficiency, compared with the 3D-HEVC. The comparison with the state-of-the-art fast algorithms further shows the superiority of the proposed algorithm.

REFERENCES

[1] Y. Chen and A. Vetro, ''Next-generation 3D formats with depth map support,'' *IEEE MultiMedia*, vol. 21, no. 2, pp. 90–94, Apr./Jun. 2014.

[2] K. Müller, P. Merkle, and T. Wiegand, ''3D video representation using depth maps,'' *Proc. IEEE*, vol. 99, no. 4, pp. 643–656, Apr. 2011.

[3] Y. Chen, M. M. Hannuksela, T. Suzuki, and S. Hattori, ''Overview of the MVC + D 3D video coding standard,'' *J. Vis. Commun. Image Represent.*, vol. 25, no. 4, pp. 679–688, May 2014.

[4] G. Tech, Y. Chen, K. Müller, 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.

[5] K. Müller *et al.*, ''3D high-efficiency video coding for multi-view video and depth data,'' *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3366–3378, Sep. 2013.

[6] Y. Chen, G. Tech, K. Wegner, and S. Yea, *Test Model 11 of 3D-HEVC and MV-HEVC*, document JCT3V-K1003, Joint Collaborative Team on 3D Video Coding Extensions (JCT-3V), 11th Meeting: Geneva, Switzerland, Feb. 2015.

[7] P. Kau *et al.*, ''Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability,'' *Signal Process., Image Commun.*, vol. 22, no. 2, pp. 217–234, Feb. 2007.

[8] C. Yan, Y. Zhang, F. Dai, X. Wang, L. Li, and Q. Dai, ''Parallel deblocking filter for HEVC on many-core processor,'' *Electron. Lett.*, vol. 50, no. 5, pp. 367–368, Feb. 2014.

[9] K. Müller, P. Merkle, G. Tech, and T. Wiegand, ''3D video coding with depth modeling modes and view synthesis optimization,'' in *Proc. IEEE Int. Conf. (APSIPAASC)*, Hollywood, CA, USA, Dec. 2012, pp. 1–4.

[10] F. Jäger, "Simplified depth map intra coding with an optional depth lookup table," in *Proc. Int. Conf. 3D Imag.*, Liège, Belgium, Dec. 2012, pp. 1–4.
[11] M. Winken, H. Schwarz, and T. Wiegand, "Motion vector inheritance for high efficiency 3D video plus depth coding," in *Proc. Picture Coding Symp.*, Krakow, Poland, May 2012, pp. 53–56.
[12] Q. Zhang, Q. Yang, Y. Chang, W. Zhang, and Y. Gan, "Fast intra mode decision for depth coding in 3D-HEVC," *Multidimensional Syst. Signal Process.*, vol. 28, no. 4, pp. 1203–1226, Oct. 2017.
[13] G. He *et al.*, "Fast mode decision and PU size decision algorithm for intra depth coding in 3D-HEVC," *J. Vis. Commun. Image Represent.*, vol. 49, pp. 303–314, Nov. 2017.
[14] Q. Zhang, N. Zhang, and T. Wei, "Fast depth map mode decision based on depth–texture correlation and edge classification for 3D-HEVC," *J. Vis. Commun. Image Represent.*, vol. 45, pp. 170–180, May 2017. doi: 10.1016/j.jvcir.2017.03.004.
[15] J. Chen *et al.*, "Sum-of-gradient based fast intra coding in 3D-HEVC for depth map sequence (SOG-FDIC)," *J. Vis. Commun. Image Represent.*, vol. 48, pp. 329–339, Oct. 2017.
[16] L. Shen, K. Li, R. Feng, and Z. Liu, "Efficient intra mode selection for depth-map coding utilizing spatiotemporal, inter-component and inter-view correlations in 3D-HEVC," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4195–4206, Sep. 2018.
[17] M. Saldanha, G. Sanchez, C. Marcon, and L. Agostini, "Block-level fast coding scheme for depth maps in three-dimensional high efficiency video coding," *J. Electron. Imag.*, vol. 27, no. 1, Feb. 2018, Art. no. 010502.
[18] G. Sanchez, L. Agostini, and C. Marcon, "A reduced computational effort mode-level scheme for 3D-HEVC depth maps intra-frame prediction," *J. Vis. Commun. Image Represent.*, vol. 54, no. 7, pp. 193–203, Jul. 2018.
[19] F. Bossen. (Oct. 2015). *3D-HEVC Software HTM-16.0*. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn3DVCSoftware/tags/HTM-16.0
[20] C. R. Wang and T. Zhang, "A fast image texture analysis algorithm," *Opto-Electron. Eng.*, vol. 32, no. 1, pp. 74–77, Jan. 2005.
[21] K. Müller and A. Vetro, *Common Test Conditions of 3DV Core Experiments*, document JCT3VG1100, Joint Collaborative Team on 3D Video Coding Extension (JCT-3V), 7th Meeting, San Jos, CA, USA, Jan. 2014.

JIABAO ZUO received the B.E. degree in communication engineering from the Changchun University of Science and Technology, Changchun, China, in 2017. She is currently pursuing the degree with the School of Information Science and Engineering, Huaqiao University, Xiamen, China. Her research interest includes video coding.

JING CHEN (M'17) received the B.S. and M.S. degrees in computer science from Huaqiao University, Xiamen, China, and the Ph.D. degree in control engineering from Xiamen University, Xiamen, China.

She was a Visiting Scholar with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, from 2011 to 2012. She is currently an Associate Professor with the School of Information Science and Engineering, Huaqiao University. Her research interests include image processing and video coding. She received the Best Paper Award from the Chinese Conference on Signal Processing 2017. She is a Traffic Video Technical Committee Member of the China Society of Image and Graphics. She was the Financial Co-Chair and the Session Chair of the IEEE International Symposium on Intelligent Signal Processing and Communication Systems 2017. She has been serving as a Reviewer for the *Journal of Visual Communication and Image Representation* and *Multimedia Tools and Applications*.

HUANQIANG ZENG (S'10–M'13–SM'18) received the B.S. and M.S. degrees from Huaqiao University, Xiamen, China, and the Ph.D. degree from Nanyang Technological University, Singapore, all in electrical engineering.

He was a Research Associate with Temasek Laboratories, Nanyang Technological University, in 2008, and a Postdoctoral Fellow with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, from 2012 to 2013. He is currently a Professor with the School of Information Science and Engineering, Huaqiao University. He has published over 80 papers in well-known international journals and conferences. His research interests include image processing, video coding, machine learning, pattern recognition, and computer vision.

Dr. Zeng is a member of the International Steering Committee of the International Symposium on Intelligent Signal Processing and Communication Systems. He received the Best Paper Award from the Chinese Conference on Signal Processing 2017. He has been serving as the Technical Program Committee Member for multiple flagship international conferences, the General Co-Chair for the IEEE International Symposium on Intelligent Signal Processing and Communication Systems 2017, the Technical Program Co-Chair for Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2017, and the Area Chair for the IEEE International Conference on Visual Communications and Image Processing 2015. He has been actively serving as an Associate Editor for *IET Electronics Letters* and the *International Journal of Image and Graphics*, and the Guest Editor for multiple international journals, including the *Journal of Visual Communication and Image Representation*, *Multimedia Tools and Applications*, the *Journal of Ambient Intelligence and Humanized Computing*.

CANHUI CAI (M'06–SM'08) received the B.S. degree from Xidian University, Xi'an, China, in 1982, the M.S. degree from Shanghai University, Shanghai, China, in 1985, and the Ph.D. degree from Tianjin University, Tianjin, China, in 2003, all in electronic engineering.

He has been with the faculty of Huaqiao University, Quanzhou, China, since 1984, where he is currently a Professor with the Engineering Institute. He was a Visiting Professor with the Delft University of Technology, Delft, The Netherlands, from 1992 to 1993, and a Visiting Professor with the University of California at Santa Barbara, Santa Barbara, CA, USA, from 1999 to 2000. He has authored or co-authored four books and published more than 120 papers in journals and conference proceedings. His research interests include video communications, image and video signal processing, and computer vision. He was a General Co-Chair of the Intelligent Signal Processing and Communication Systems, in 2007.

KAI-KUANG MA (S'80–M'84–SM'95–F'13) received the B.E. degree in electronic engineering from Chung Yuan Christian University, Taoyuan, Taiwan, the M.S. degree in electrical engineering from Duke University, Durham, NC, USA, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, NC, USA.

From 1984 to 1992, he was with IBM Corporation at Kingston, Kingston, NY, USA, and Research Triangle Park, NC, USA, focusing on various DSP and VLSI advanced product development. From 1992 to 1995, he was a member of Technical Staff with the Institute of Microelectronics, Singapore, where he was involved in digital video coding and the MPEG standards. He is currently a Full Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He has published extensively and holds one USA patent on fast motion estimation algorithm. His research interests include digital image/video processing and computer vision, including digital image/video coding and standards, image/video segmentation, denoising, enhancement, interpolation, super resolution, image matching and registration, scene analysis and recognition, and human–computer interaction.

Dr. Ma is a member of Sigma Xi and Eta Kappa Nu. He is an elected member of three IEEE Technical Committees: The Image and Multidimensional Signal Processing Committee, the Multimedia Communications Committee, and the Digital Signal Processing. He was serving as the Singapore MPEG Chairman and the Head of Delegation, from 1997 to 2001. On the MPEG contributions , two fast motion estimation algorithms (Diamond Search and MVFAST) produced from his research group have been adopted by the MPEG-4 standard, as the reference core technology for fast motion estimation. He was the General Chair of organizing a series of international standard meetings (MPEG and JPEG), JPEG2000, and MPEG-7 workshops held in Singapore, in 2001. He has been serving as the Technical Program Committee Member, Reviewer, and the Session Chair of multiple IEEE international conferences. He is the General Co-Chair of ISPACS2017, ASIPA2017, ACCV2016 Workshop, and VCIP-2013, the Technical Program Co-Chair of ICIP-2004, ISPACS-2007, IIH-MSP-2009, and PSIVT-2010, and the Area Chair of ACCV-2009 and ACCV-2010. He is the Chairman of the IEEE Signal Processing Singapore Chapter, from 2000 to 2002. He was elected as a Distinguished Lecturer of the IEEE Circuits and Systems Society, for the period 2008–2009. He has been serving as an Editorial Board Member for several leading international journals in his research area, such as a Senior Area Editor for the IEEE Transactions on Image Processing, from 2016 to 2019, an Editor for the IEEE Transactions on Communications, from 1997 to 2012, and an Associate Editor for the IEEE Transactions on Multimedia, from 2002 to 2009, the *International Journal of Image and Graphics*, from 2003 to 2015, the *Journal of Visual Communication and Image Representation*, from 2005 to 2015, the IEEE Transactions on Image Processing, from 2007 to 2010, the IEEE Signal Processing Letters, from 2014 to 2016, and the IEEE Transactions on Circuits and Systems for Video Technology, since 2015.