Fast Intra Mode Decision for High Efficiency Video Coding (HEVC)

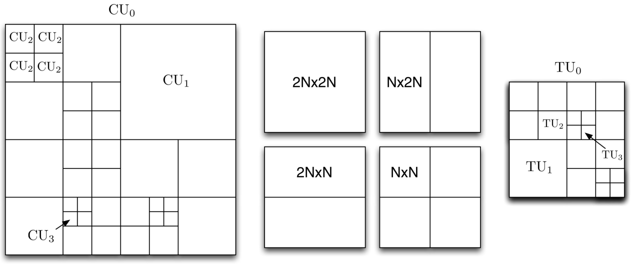
Hao Zhang, *Member, IEEE,* and Zhan Ma, *Member, IEEE*

***Abstract*—The latest High Efficiency Video Coding (HEVC) standard only requires 50% bit-rate of the H.264/AVC at the same perceptual quality, but with a significant encoder complexity increase. Hence, it is necessary and inevitable to develop fast HEVC encoding algorithms for its potential market adoption. In this paper, we propose a fast intra mode decision for the HEVC encoder. The overall fast intra mode decision algorithm consists of both micro- and macro-level schemes. At the micro-level, we propose the Hadamard cost-based progressive rough mode search (pRMS) to selectively check the potential modes instead of traversing all candidates (i.e., up to 35 in HEVC). Fewer effective candidates will be chosen by the pRMS for the subsequent rate-distortion optimized quantization (RDOQ) to derive the rate-distortion (R-D) optimal mode. An early RDOQ skip method is also introduced to further the complexity reduction. At the macro-level, we introduce the early coding unit (CU) split termination if the estimated R-D cost [through aggregated R-D costs of (partial) sub-CUs] is already larger than the R-D cost of the current CU. On average, the proposed fast intra mode decision provides about 2.5 × speedup (without any platform or source code level optimization) with just a 1.0% Bjontegaard delta rate (BD-rate) increase using the HEVC common test condition. Moreover, our proposed solution also demonstrates the state-of-the-art performance in comparison with other works.**

***Index Terms*—High Efficiency Video Coding (HEVC), intra prediction, mode decision, video coding.**

I. INTRODUCTION

THE LATEST video coding standard, High Efficiency Video Coding (HEVC), which was developed under the efforts of the Joint Collaborative Team on Video Coding (JCT-VC), has demonstrated a significant performance improvement on compression ratio over the H.264/AVC [1]–[3]. It is noted that part of the coding efficiency improvement comes from the newly introduced recursive quad-tree-based coding unit split (i.e., from 64×64 to 8×8) [4], larger block transforms (i.e., up to 32×32), fine-grain spatial intra prediction [5], etc. Macroblock in the previous standards is extended in HEVC to the coding unit (CU), prediction unit (PU), and transform unit (TU). Each CU can be recursively split into four sub-CUs and at each CU (or sub-CU) level, it can be further split into one or multiple PUs. Moreover, HEVC also adopts recursive TUs at each PU level for residual coding. Fig. 1 gives a brief illustration of the recursive CU, PU, and TU in HEVC. More details can be found in [1], [2], and [4].



(a) (b) (c)

Fig. 1. Recursive block unit structure in HEVC, where k indicates the depth for CU_k and TU_k. (a) CU structure. (b) PU structure. (c) TU structure.

Spatial intra prediction [5] has been significantly improved in HEVC over H.264/AVC with more fine-granular predictions, where up to 35 prediction modes (i.e., mode 0 for PLANAR, mode 1 for DC, and others are angular modes) are used in HEVC versus up to nine modes in H.264/AVC. As we can see, the increased number of spatial intra prediction modes leads to much higher compression efficiency, while requiring tremendous computational power using the expensive rate-distortion optimization (RDO) to search for the optimal mode in a brute-force fashion. There are two RDO methods implemented in the HEVC reference software HM [6]. The default one is the high-efficiency rate-distortion optimized quantization (RDOQ) based RDO, in which a soft decision quantization is performed for a coefficient given both its impact on the bit-rate and quality [7]. For convenience, we simply refer to this RDOQ based RDO method as RDOQ in this paper. The other method applies the uniform quantization (or hard decision quantization) for every coefficient, which incurs lower complexity compared with the RDOQ but also comes with noticeable coding efficiency loss.

Given the unbearable complexity to choose the optimal intra prediction mode through brute-force RDOQ, HEVC reference software adopts a four-step method,[1] in which rough mode decision (RMD) is first performed using the Hadamard cost ranking to choose fewer candidates (out of 35) [8]. Then,

Manuscript received June 25, 2013; revised August 27, 2013; accepted September 24, 2013. Date of publication November 12, 2013; date of current version April 2, 2014. This paper was recommended by Associate Editor B. Pesquet-Popescu.

H. Zhang is with the School of Information Science and Engineering, Central South University, Changsha 410083, China (e-mail: hao@csu.edu.cn).

Z. Ma is with Samsung Research America at Dallas, Richardson, TX 75082 USA (e-mail: zhan.ma@ieee.org).

Digital Object Identifier 10.1109/TCSVT.2013.2290578

---

[1] It was a three-step method originally proposed in [8] and later refined with four-step with improved performance by adding the most probable modes selection.

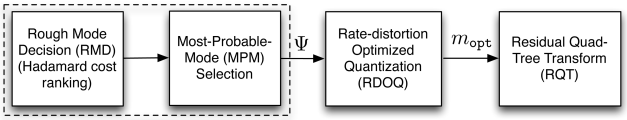


Fig. 2. intra mode decision process in HEVC reference software. Ψ is the effective candidate set after performing the Hadamard cost ranking and MPM selection; $m_{\texttt{opt}}$ is the best mode chosen by the RDOQ from Ψ for the subsequent RQT.

additional most probable modes (MPM) will be added to these few candidates if they are not included yet, given the assumption that MPM has the high probability to be chosen because of the high correlation of spatial neighbors. RDOQ is then applied to these effective candidates at the maximum TU size. In the end, for the best mode chosen from the RDOQ, residual quadtree-based TU selection (RQT) is further utilized to derive the optimal TU size. Fig. 2 pictures the main idea of such a four-step method for the optimal prediction mode decision at each CU depth. Moreover, HEVC reference encoder implementation [6] goes through all the CU depths recursively from the largest CU (LCU) and finally reaches the stage with the optimal CU size and optimal prediction mode. Even though the four-step-based intra mode decision speedup is applied [8], it still requires massive computing resources to evaluate each potential mode at each possible CU depth level. Hence, a fast encoder implementation is extremely desirable for the practical HEVC-based encoding applications.

Following the discussion above, the overall intra mode decision algorithm involves optimal prediction mode decision at each CU level, and optimal CU size (or split depth) decision. To formulate the problem well, we note the CU level intra prediction decision processes as the micro-level operation, and the CU size decision as the macro-level operation. Thus, in this paper, we develop a novel fast intra mode decision for HEVC, which consists of fast algorithms for both micro- and macro-level processes. Briefly, at micro-level, we propose the progressive rough mode search (pRMS) to selectively go through the available modes (rather than traversing all of them in the default rough mode decision [8]) and the early RDOQ skip scheme to further remove unlikely modes with more complexity reduction; on the other hand, an early CU split termination method is utilized to avoid unnecessary computing based on the aggregated rate-distortion (R-D) cost (of partial sub-CUs) at macro-level. Experiments report the state-of-the-art performance in terms of the tradeoff between Bjontegaard delta rate (BD-rate) loss and encoder complexity reduction for our proposed algorithm. More detailed simulation explanation will be shown in Section IV.

The remainder of this paper is organized as follows. Section II presents a brief literature review of the fast intra prediction algorithms for HEVC. Section III describes the proposed fast intra mode decision of HEVC at both macro and micro perspectives. Extensive experiments are carried out and analyzed in Section IV to demonstrate the efficiency of the proposed solution, while Section V concludes the work in this paper.

II. FAST INTRA MODE DECISION FOR HEVC: A REVIEW

Fast intra mode decision has been extensively studied for the H.264/AVC, with few of them mentioned here as an example [9]–[15]. However, these algorithms cannot be directly applied to HEVC due to quite different coding structures and prediction modes. Recently, a few fast intra mode decision algorithms have been proposed for HEVC (some algorithms are developed for both inter decision and intra decision). In general, these algorithms could be categorized into three main classes, i.e., fast CU or PU size decision, fast intra prediction mode decision and fast RQT, which are briefly summarized as follows.

A. *Fast Coding Unit Size Decision*

Choi and Jang [16] have developed a tree pruning algorithm for fast CU size decision, based on the observation that if the current CU chooses the SKIP mode as the best mode, then no further splitting is required. This algorithm is implemented and reported on HM4.0, with almost 40% encoding time reduction. However, this method is applied to the inter mode only.

A gradient based fast intra mode decision was proposed in [17], where gradient directions and histogram are derived for CU size decision. With this approach, 20% time saving, on average, is achieved with negligible loss of coding efficiency on HM4.0.

Shen *et al.* [18] recently proposed a fast mode decision scheme based on the Bayesian decision rule. The split and nonsplit decision is made on the Bayesian risk, which can be calculated from the Lagrangian cost, the class-conditional probability density functions, and priori probabilities. The feature vectors are calculated online, while other parameters are calculated off-line. Random access (RA) and low delay (LD) configurations are used for simulations and, on average, 41.4% encoding time saving is reported.

Tian and Goto [19] proposed a two-stage fast intra mode decision based on the texture complexity analysis. In the prestage, LCUs are down-sampled to 16x16 blocks to derive the texture complexity. The intuition is that when the complexity is less than a certain threshold, small PUs are unlikely to be used and vice versa. The proposed algorithm achieves 44.9% encoding time reduction for 4kx2k sequences and 28.8% for 1080p videos on average.

Shen *et al.* [20] proposed early termination schemes to speed up the mode decision according to the spatial correlation between neighbor CUs. However, many techniques in this paper such as early termination of motion estimation are applied for inter mode rather intra mode decision.

A piece of our earlier work includes a set of early termination techniques for fast intra mode decision [21] based on the experimental observations. The R-D cost of four sub-CUs and sub-TUs is predicted by the Hadamard costs. The predicted cost is then compared with the R-D cost in the upper level to decide whether the CU split is necessary. As will be shown in Section III-C, we have updated the algorithm significantly with in-depth analytical discussions on why the early termination scheme works.

B. Fast Intra prediction Mode Decision

Zhao *et al.* [22] have studied the impact of the number of mode candidates after the RMD process in HM. The proposed method basically tried to reduce the number of candidates for the full RDOQ process. It achieved averaged 20% and 28% encoding time saving in high efficiency (HE) and low complexity (LC) test conditions on HM1.0, with a negligible loss of compression efficiency. It is noted that HE and LC test settings were used in the early phase of HEVC development, which were merged later then.

Zhang and Ma [23] have proposed a three-step solution to speed up the HEVC intracoding. At the RMD step, the original Hadamard transform is replaced by a 2:1 down-sampled Hadamard transform to derive the cost for evaluation, followed by a gradual progressive mode search. Meanwhile, an early termination method is also included to reduce the number of modes for the subsequent RDOQ. It reports averaged 38% encoding time reduction with 2.9% BD-rate loss of the luma component on HM6.0 using all JCT-VC test sequences.

C. Fast RQT

In HEVC, RQT is applied to further exploit the spatial correlation and improve the coding efficiency [1] by implementing recursive TU splitting. It demands significant computational overhead. Hence, Tan *et al.* [24] proposed fast RQT algorithms for both intra mode and inter mode coding so as to reduce the complexity. As reported using HM2.0, for all intra (AI) case, fast RQT saves 13% encoding time with 0.1% BD-rate increase. For RA and LD scenarios, the proposed fast RQT algorithm reduces up to 9% encoding time at the expense of (up to) 0.3% BD-rate performance degradation.

Another RQT coding scheme was proposed in [25] that replaces the original depth-first mode decision process with a merge-and-split decision process. They proposed early termination of the merge-and-split processes, i.e., when the current TU is a zero-block, no further split is performed any more. It reports almost 2 × speedup for random access using the HE configuration (for class C and D non-HD sequences only) on HM2.0, with about 0.3% BD-rate increase.

Furthermore, Zhang and Zhao [26] have developed the adaptive RQT algorithms for inter mode by restricting the smaller RQT depth for larger CU size and vice verse. It shows 0.7% BD-rate increase with 7.2% encoder complexity reduction for HE configuration, and 0.6% BD-rate increase with almost 21% encoder complexity reduction for LC configuration on HM4.0.

III. PROPOSED FAST INTRA PREDICTION ALGORITHM

As briefed in Section I, spatial intra prediction goes through every possible CU size recursively, and at each CU level, a tremendous amount of prediction modes (i.e., up to 35 for intraluma component) are tested to derive the R-D optimal one. Then, the speedup of the overall process could be roughly categorized into two main parts: one is early terminating the check of unnecessary CU size which obviously will not be chosen as the optima and the other one is quickly determining the most effective prediction candidates set (which has less modes) without going through all modes in a brute-force way at a certain CU level. We note the former one as the macro operations and the latter one as the micro operations, both for the intra prediction speedup. As shown in the following discussions, some of the intermediate data produced by the micro level could be used by the macro-level scheme, we organize the overall algorithm description from the micro to the macro perspective for convenience.

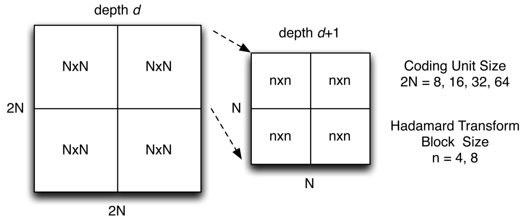


Fig. 3. Illustration of 2N × 2N CU, four N × N sub-CUs, and Hadamard transform block.

A. Brief of Hadamard Cost Derivation

In addition to the normal rate-distortion cost (noted as J^{RD}) derived from the expensive RDOQ, Hadamard cost (defined as Hcost or J^{H}) is utilized to reduce the complexity for encoder-side mode decision. This Hcost can be calculated by deriving the sum of the absolute Hadamard transformed difference and the estimated mode bits consumption. It is adopted in the HEVC reference software HM [6] to speed up the encoding process, and is directly applied in our proposed schmes without any change.

As shown in Fig. 3, for a certain 2Nx2N CU (and its four N×N sub-CUs), each nonoverlapped 8×8 or 4×4 block (depending on whether 2N > 8) will produce its individual Hcost, and then the total Hcost for any CU is the sum of the all block wise Hcost within this CU. More details regarding the implementation of the Hadamard cost calculation in HM could be found in [6] and [8].

B. Micro-Level Algorithm for Fast Prediction Modes Decision

1) *Progressive Rough Mode Search (pRMS):* For existing fast motion estimation (FME) algorithms, such as the well known three-step or diamond search [27], [28], they typically involve several steps. At each step, its optimal (in the sense of the lowest rate-distortion cost) location is selected as the new search center for the next step. Also, the search range gets smaller step by step before finally converging to the global best location. Enlightened by the FME algorithms, we propose pRMS for the rough mode decision in HEVC.

Before jumping into details, we first use the difference between intra mode indices to define the mode distance. For instance, mode i is a d-distance neighbor of mode j if $|i - j| = d, 2 \leq i, j \leq 34$. PLANAR ($i = 0$) and DC ($i = 1$) modes are treated separately since they are not the same as the other angular modes. In this paper, we simply assume that they do not have neighbors.

During the rough mode decision process, instead of traversing all modes, we propose to check equally spaced four-distance nine modes plus PLANAR and DC modes, i.e., 0, 1, $2 + 4 \cdot \delta$, $0 \leq \delta \leq 8$, and to order them along with the Hcost increase monotonically (noted as Λ). For the best six modes (i.e., with least four Hcosts) in Λ, we then check their two-distance neighbors, and encoded modes of the above and left CUs if they are not tested yet. Subsequently, the one-distance neighbors are evaluated for the best two modes from the previous step. Finally, MPMs are added to Λ if they are not checked. Here, the MPM selection is the same as the default RMD without change. Please note that, at each step, we update Λ by adding possible new modes (if applicable) and maintain the order where the associated Hcost increases monotonically. As we can see, this pRMS is very like the FME where the search range gradually decreases. We then pick up M modes from the final Λ for the full RDOQ process. Please note that the number of candidates (i.e., M) for the full RDOQ process is the same for our proposed pRMS and the default rough mode decision in HM [6], i.e., M=$\{8, 8, 3, 3, 3\}$ for CU size at $\{4\times4, 8\times8, 16\times16, 32\times32, 64\times64\}$, respectively.

In summary, pRMS first decomposes the total 35 modes into 10 groups, i.e., PLANAR and DC, and equally four-distance spaced angular modes. This could be seen as the coarse directional decomposition, which can capture the main angle of the texture. Few modes are chosen to evaluate their two-distance and further one-distance neighbors based on the Hcost ranking step-by-step. This can be referred as the fine grain refinement. To well understand the proposed pRMS, we then show an illustrative example as follows.

1) Initially, $\{0, 1, 2, 6, 10, 14, 18, 22, 26, 30, 34\}$ are selected and checked with the best six modes being $\{0, 6, 10, 14, 18, 26\}$.
2) The two-distance neighbors of mode 6, 10, 14, 18, 26 are $\{4, 8\}$, $\{8, 12\}$, $\{12, 16\}$, $\{16, 20\}$, $\{24, 28\}$, respectively. Additional modes introduced by the two-distance neighbors are $\{4, 8, 12, 16, 20, 24, 28\}$. Meanwhile, we also include the modes from the spatial neighbors of current CU, i.e., upper and left. Let their modes be $\{4, 8\}$. Then, Λ is updated after checking the modes $\{4, 8, 12, 16, 20, 24, 28\}$.
3) Let the best two modes be the $\{4, 6\}$ from the latest updated Λ. The one-distance neighbors of mode 4 and 6 are respective $\{3, 5\}$ and $\{5, 7\}$. Therefore modes $\{3, 5, 7\}$ are evaluated and put into the Λ appropriately.
4) MPMs of current CU are added into the ordered Λ if they are not checked yet.
5) Select the best M candidates from ordered Λ to form the effective candidate set Ψ for subsequent full RDOQ process.

As aforementioned, both pRMS and default RMD will produce the ordered effective candidate set Ψ for later RDOQ process. It is also proven in [22] that it has very high probability that the best mode chosen by the RDOQ is coming from the first two or three modes in Ψ (i.e., corresponding to the two or three least Hcost). Therefore, to verify the efficiency of the proposed pRMS, we collect the modes from both pRMS and default RMD, and show the percentage that first three modes are the same for both schemes in Table I. The first three modes produced by the pRMS is highly correlated with the results from the default RMD, which indicates the potential effectiveness of our proposed method.

TABLE I
PERCENTAGE OF THE SAME FIRST THREE MODES DERIVED BY THE pRMS AND THE DEFAULT RMD

Sequences	QP	CU depth d				
		0	1	2	3	4
BasketballPass	22	66.8%	76.4%	84.9%	99.7%	99.6%
QWVGA	37	69.8%	82.9%	85.8%	99.6%	99.8%
Soccer	22	74.0%	71.1%	81.1%	99.7%	99.6%
4CIF	37	81.5%	77.3%	82.1%	99.5%	99.7%
Vidyo1	22	89.7%	87.9%	90.2%	99.4%	99.5%
720p	37	90.0%	88.1%	89.7%	99.6%	99.8%

Meanwhile, Table I also indicates that the correlation between pRMS and default RMD is getting larger at smaller size CU (or larger CU depth). Typically, a larger size CU contains richer content distribution (such as texture, edge), and we also assume that the modes close to the dominant direction will have smaller Hcost. For instance, a 64×64 CU may contain two objects with different dominant spatial directions. Default RMD produces modes $\{2, 9, 5\}$ as the best three modes, with mode 2 and 9 as two distinct directions. However, pRMS will produce modes $\{2, 6, 10, 14, 18, 26, 30\}$ for the two-distance modes search. Mode 2 and mode 10 are filtered out to check their one-distance neighbors (i.e., mode 3, 9 and 11), resulting in the final best three modes $\{2, 9, 10\}$ or $\{2, 9, 3\}$ or $\{2, 9, 11\}$, etc. We can see that mode 5 is missed by pRMS utilizing the coarse-to-fine selection procedure rather than the exhaustive search in default RMD. On the other hand, a smaller size CU would have more stationary content distribution, with higher probability to have single dominant spatial direction. In this case, pRMS will potentially have larger chance to select the same modes as the default RMD. Reusing the aforementioned example for a 4×4 CU, default RMD will select $\{4, 3, 5\}$ as the best three modes, with mode 4 as the dominant edge direction. pRMS will have the same modes $\{2, 6, 10, 14, 18, 26, 30\}$ for their two-distance modes search. Modes $\{4, 2\}$ are then chosen as the best two candidates to evaluate their one-distance neighbors (i.e., mode 3 and 5). Finally, modes $\{4, 3, 5\}$ are selected as the best three modes which are the same as the default RMD. As will be further shown in later experiments (Prop.-pRMS in Table VII), pRMS can efficiently reduce the encoder complexity with negligible performance loss.

2) *Early RDOQ Skip:* As demonstrated above, pRMS effectively reduces the complexity of the original default RMD in the HEVC reference software without looping over up to 35 modes. However, we still keep the number of effective candidates produced by the pRMS the same as the default RMD. These modes will be further evaluated through the RDOQ and later RQT[2] to derive the best one in a R-D optimal manner.

It is expected that the complexity of the micro-level decision could be further reduced if we can skip some modes from

---

[2] For convenience, we simply use the term RDOQ to represent the overall process of RDOQ based RDO and subsequent RQT (where RDOQ based RDO is also involved to decide the best TU).

Algorithm 1 Algorithm for early skip of the RDOQ

```
//pick $M$ effective modes from pRMS to form $\Psi$
$S = \{m_1, m_2\}$
for  $3 \leq i \leq M$ do
    if $m_i$ is a near-neighbor of any mode in $S$ then
        //skip $m_i$ for the subsequent RDOQ
    else
        $S = S + m_i$
    end if
    i++
    //let $S_0 = \{m_1, m_2, \mathrm{PLANAR}, \mathrm{DC}, \mathrm{MPM}\}$
    if  $S_0 \subset S$  then
        break
    end if
end for
```

the effective candidate set produced by the pRMS. Hence, we propose an early RDOQ skip scheme based on the mode adjacency and associated Hcost (which was calculated already).

Let Ψ be the effective candidate set after the pRMS, i.e., $\Psi = \{m_1, m_1, ..., m_M\}$. M Modes are ordered along with their Hcosts from the lowest to the highest. The final best mode $m_{\mathtt{opt}}$ would be chosen from these M modes from the subsequent RDOQ process. Meanwhile, as shown in [22], the first few (i.e., two or three) mode candidates (corresponding the lowest and the second lowest Hcost) in Ψ have the high probability to be selected as the optimal mode $m_{\mathtt{opt}}$. Here, first two modes are chosen to always go through the full RDOQ check in our scheme. For an arbitrary mode $m_i \in \Psi$ except m_1 and m_2, we first check whether it is a near-neighbor of any mode evaluated already. If it is a near-neighbor, m_i will be excluded for RDOQ process; otherwise it is added for subsequent RDOQ. For simplicity, we define the relationship of the near-neighbor if their mode difference is less than 2. Furthermore, as long as we have tested m_1, m_2, PLANAR, DC and MPM,[3] we will skip all the rest modes in Ψ.

Following the previous definitions and constraints, early skip of the RDOQ (and later RQT) could be summarized in Algorithm 1. We select several different video sequences (from QWVGA to 720p), and encode them using QPs at 22, 27, 32, and 37 with all intrasettings. We then collect the hit ratio η of the optimal mode, which is calculated by comparing whether the results are the same for the default RDOQ mode selection process and the proposed solution with early RDOQ skip; and the percentage of the RDOQ reduction which is approximated by the $\Delta\mathrm{RDOQ} = \left(M - M_{\mathtt{es\_RDOQ}}\right)/M \times 100\%$ with M and $M_{\mathtt{es\_RDOQ}}$ indicating the number of modes, again for the default RDOQ mode selection process and the proposed solution with early RDOQ skip, respectively. Table II shows that early RDOQ skip can efficiently reduce the complexity by removing the unlikely modes while preserving the performance, i.e., nearly half of the modes are early skipped with on average $\approx 95\%$ accuracy for the optimal mode decision.

TABLE II

STATISTICAL DISTRIBUTION (HIT RATIO η AND RDOQ REDUCTION ΔRDOQ) OF EARLY RDOQ SKIP FOR DIFFERENT VIDEOS

Sequence	Statistics	QP			
		22	27	32	37
BasketballPass QWVGA	η ΔRDOQ	91.22% 47.94%	94.26% 50.30%	96.60% 53.27%	98.26% 56.46%
Soceer 4CIF	η ΔRDOQ	89.49% 47.26%	92.25% 49.52%	94.89% 52.65%	97.34% 56.22%
Mobisode2 WVGA	η ΔRDOQ	97.35% 55.93%	98.52% 58.03%	99.22% 59.48%	99.62% 60.86%
OldTownCross 720p	η ΔRDOQ	89.53% 47.20%	93.50% 50.93%	95.73% 53.45%	97.66% 56.12%

C. Macro-Level Algorithm for Early Coding Unit Split Termination Using the Aggregated R-D Cost of Partial Sub-CUs

In addition to the pRMS and early RDOQ skip as the micro-level scheme, our proposed method also employs the Hcosts (before the actual R-D optimized encoding) and aggregated actual R-D costs of partial sub-CUs to further accelerate the intra mode decision process by early terminating the CU split.

Typically, the R-D cost at the current depth level d (referred to as $J_d^{\mathtt{RD}}$) is always compared with the sum of the R-D costs of its four NxN sub-CUs (noted as $J_{d+1}^{\mathtt{RD}} = \sum_{i=1}^{4} J_{d+1}^{\mathtt{RD}}(i)$). The CU depth level with the smaller R-D cost is selected for the next step. These four sub-CUs are encoded in a raster-scan order, i.e., from the left-top sub-CU to the right-bottom sub-CU. We use $J_{d+1}^{\mathtt{RD}}(i)$ to represent the R-D cost for the ith sub-CU ($1 \leq i \leq 4$) and use $J_{d+1,K}^{\mathtt{RD}}$ to denote the aggregated R-D cost after encoding the K ($K \in [1, 2, 3, 4]$) NxN sub-CUs (given then 2Nx2N CU). Hence, we have

$$J_{d+1,K}^{\mathtt{RD}} = \sum\nolimits_{i=1}^{K} J_{d+1}^{\mathtt{RD}}(i). \tag{1}$$

It requires the actual RDOQ process to finish all sub-CU encoding (i.e., $K = 4$) and obtain the total R-D cost with the split option. This cost will be compared with the R-D cost without split, and the one with the less cost will be chosen for the next step. As we can see, the complexity could be reduced if we can estimate the total R-D cost with the split option without finishing all sub-CU encoding (i.e., $K < 4$). Toward this purpose, we propose to use the predicted R-D cost $\hat{J}_{d+1}^{\mathtt{RD}}$ as the approximation of the $J_{d+1}^{\mathtt{RD}}$ ($J_{d+1}^{\mathtt{RD}} = J_{d+1,K=4}^{\mathtt{RD}}$) using the aggregated R-D costs after completing the encoding of partial sub-CUs. The RDOQ based on-going split process will be early terminated if the following inequality holds, that is

$$\hat{J}_{d+1}^{\mathtt{RD}} > \beta_K J_d^{\mathtt{RD}}. \tag{2}$$

To derive the predicted R-D cost, we first use the $J_{d+1,K}^{\mathtt{H}}$ to represent the aggregated Hcost at depth level $d + 1$, and then $\hat{J}_{d+1}^{\mathtt{RD}}$ is approximated via

$$\hat{J}_{d+1}^{\mathtt{RD}} = \left(\min\left\{\frac{4}{K}, \frac{J_{d+1,K=4}^{\mathtt{H}}}{J_{d+1,K}^{\mathtt{H}}}\right\}\right) \cdot J_{d+1,K}^{\mathtt{RD}}. \tag{3}$$

Here, we assume that the ratio of the K aggregated Hcost versus the total Hcost is highly correlated with the ratio of the K aggregated R-D costs versus the total R-D costs. Note that $J_{d+1,K=4}^{\mathtt{H}}$ or $J_{d+1}^{\mathtt{H}}$ denotes the total Hcost for all four NxN

---

[3] It is noted that these modes will be overlapped, i.e., m_1 might be the same as DC; MPM could be PLANAR, etc.

sub-CUs at level $d+1$ (given the 2Nx2N CU at depth d). It is also noted that $J^{\mathrm{RD}}_{d+1,K}$ is the aggregated actual R-D costs after finishing the encoding of K sub-CUs.

Parameters β_K with $K = 1, 2, 3, 4$ can be seen as the prediction confidence scaling factor. For instance, in the case of $K = 4$, four sub-CUs are encoded with complete R-D cost, i.e., $\hat{J}^{\mathrm{RD}}_{d+1} = J^{\mathrm{RD}}_{d+1,K=4}$, where actual R-D cost is used to decide whether or not the current CU is required to be split. Hence, we are fully confident of reaching the R-D optimal decision without requiring any scaling, i.e., $\beta_4 = 1$. On the other hand, by allowing $\beta_K > 1$ when $K < 4$, we can efficiently reduce the false negative prediction of CU split decision, but may also miss the true positive prediction. Nevertheless, we prefer such a conservative strategy with more confidence by giving the correct CU split decision. Furthermore, larger K usually comes with more information/knowledge from fully R-D coded sub-CUs, resulting in more confidence of decision making. Therefore, β_K is a monotonically decreasing function as K increases till reaching $\beta_4 = 1$. In this paper, we encode few video sequences (*BasketballPass* at QWVGA, *Soccer* at 4CIF, *Vidyo1* at 720p) with QPs at 22 and 37, and collect the aggregated Hcost, R-D cost and ground-truth CU split decision (i.e., via actual R-D optimized recursively search) at each CU level. We then predict the CU split using the aggregated Hcost, R-D cost, and β_K via (4), and calculate the prediction accuracy by comparing the CU spit prediction with the ground-truth decision. β_Ks are adapted empirically, and finally $\{\beta_1, \beta_2, \beta_3, \beta_4\} = \{1.5, 1.2, 1.1, 1\}$ is chosen with the averaged percentage of true positive prediction of CU split over 90% for above training sequences. Then, we simply apply these fixed β_Ks for all tests, and the following simulations demonstrate that these empirically derived β_K can produce a decent balance between coding efficiency and complexity reduction for HEVC encoding. We have also tried QP adaptive β_K; however, the performance difference compared with the fixed values is marginal.

Combining (1)–(3), we reach

$$\left(\min\left\{\frac{4}{K}, \frac{J^{\mathrm{H}}_{d+1,K=4}}{J^{\mathrm{H}}_{d+1,K}}\right\}\right)\sum_{i=1}^{K} J^{\mathrm{RD}}_{d+1}(i) > \beta_K J^{\mathrm{RD}}_{d}. \qquad (4)$$

If (4) holds for any K, CU split is terminated. For a dth level 2Nx2N CU, we might only need to encode the first $(d+1)$-th level NxN sub-CU, i.e., $K = 1$, to early terminate the CU split decision without requiring the encoding of the subsequent sub-CUs. There is also the case that we have to encode all sub-CUs before reaching the final split-or-not decision, i.e., $K = 4$. In such a case, the complexity is the same as the normal RDOQ process.

Hypothetically, if $K < 4$ occupies a noticeable percentage in practical scenarios, it will result in the nonnegligible complexity reduction by applying the aggregated R-D cost (of partial sub-CUs) based early CU split termination. To confirm our assumption, we encode different videos using different QPs at 22, 27, 32, 37, and collect the statistics of $K < 4$ at different CU depths (i.e., $d = 0, 1, 2$). Results are shown in Table III, where we can explicitly see the noticeable percentage of $K < 4$, which implies that the potential encoding complexity reduction is due to the early CU split termination, as demonstrated in (4).

TABLE III

STATISTICAL DISTRIBUTION OF $K < 4$

Sequences	d	QP 22	QP 27	QP 32	QP 37
BasketballPass QWVGA	0	0.82%	3.32%	6.5%	11.71%
	1	12.6%	20.6%	29.23%	37.33%
	2	29.63%	43.36%	52.44%	61.43%
PartyScene WVGA	0	1.49%	2.31%	3.32%	5.21%
	1	1.14%	3.18%	5.6%	10.02%
	2	4.84%	9.79%	17.15%	27.04%
RaceHorses WVGA	0	8.02 %	10.71%	12.74%	14.7%
	1	10.84%	16.57%	22.31%	28.36%
	2	23.57%	32.87%	39.09%	45.88%
Vidyo1 720p	0	6.96%	9.95%	13.05%	19%
	1	9.79%	23.64%	34.69%	47.05%
	2	35.6%	53.86%	63.82%	70.99%

Moreover, for the same sequence, it shows that the percentage of $K < 4$ is larger at higher QPs (i.e., QP 37 versus QP 27), which indicates that more blocks will use the larger block size at higher QPs. This is due to the heavy quantization at larger QP where most of the local rich texture will be removed, and hence larger block size is preferred. Meanwhile, we can see that the percentage of $K < 4$ is increasing along with the increase of the CU split depth. This is because large coding unit (i.e., 64×64 LCU) typically contains more details (such as multiple objects), and it is more efficient to encode it at next depth after split. Also, along with the CU split depth increases, i.e., such as 16×16 CU, it might contain only one single object with more structured texture, which can be encoded efficiently without further block split. On the other hand, different content has the different statistical distribution of the $K < 4$. Note that *PartyScene* sequence contains more complex texture in the source in comparison to the *Vidyo1* or *RaceHorses* where most of the background is quite stationary. Therefore, *PartyScene* prefers smaller CU (i.e., more block split with lower percentage of $K < 4$) than *Vidyo1* or *RaceHorses*.

IV. PERFORMANCE VERIFICATION AND COMPARISON

A. Performance Verification

Experimental results for various test sequences are shown in this section. Our proposed fast intramethod is compared with the default algorithm in HM10 main profile, strictly following the common test conditions defined in [29] unless we point out otherwise explicitly. As shown later, we change the QPs to test the performance at both low and high bit-rate scenarios, but other settings are kept without any modification. All intraencoder setting is simulated to demonstrate encoder performance. Class A (4Kx2K), B (1080p), C (WVGA), D (QWVGA), and E (720p) sequences are all used for performance verification. BD-rate performance and encoder time reduction ΔT are shown in Table IV (noted as Prop.-Overall), where coding efficiency loss is measured using BD-rate [30] and encoder time reduction is derived by $\Delta T = \left(T_{\mathtt{org}} - T_{\mathtt{prop}}\right)/T_{\mathtt{org}} \times 100\%$ with $T_{\mathtt{org}}$ and $T_{\mathtt{prop}}$ for the encoding time of the original intra mode decision and the

TABLE IV

CODING EFFICIENCY AND COMPLEXITY REDUCTION FOR PROPOSED FAST INTRA PREDICTION ON HM10 MAIN PROFILE USING HEVC COMMON TEST CONDITION [29] WITH QPS AT 22, 27, 32, AND 37

Class & Sequences		Luma-Y	Chroma-U	Chroma-V	ΔT
A	Traffic	1.2%	0.6%	0.1%	60%
	PeopleOnStreet	1.1%	1.0%	0.8%	58%
	Nebuta	0.2%	0.5%	0.4%	61%
	SteamLocomotive	0.6%	1.6%	1.2%	67%
B	Kimono	0.6%	0.8%	0.6%	66%
	ParkScene	0.6%	0.8%	0.7%	59%
	Cactus	1.2%	0.4%	0.8%	59%
	BasketballDrive	1.3%	1.4%	1.4%	62%
	BQTerrace	1.0%	2.4%	2.8%	59%
C	BasketballDrill	1.5%	2.4%	2.4%	56%
	BQMall	1.1%	1.3%	1.4%	58%
	PartyScene	0.8%	1.3%	1.5%	53%
	RaceHorses	0.9%	1.0%	1.4%	59%
D	BasketballPass	1.2%	1.5%	1.6%	58%
	BQSquare	0.9%	1.2%	1.5%	53%
	BlowingBubbles	0.9%	1.5%	1.8%	52%
	RaceHorses	1.0%	1.7%	1.8%	56%
E	FourPeople	1.2%	1.5%	1.4%	61%
	Johnny	1.8%	2.6%	3.0%	66%
	KristenAndSara	1.6%	2.2%	2.3%	65%
Ave.		1.0%	1.4%	1.4%	60%

TABLE V

CODING EFFICIENCY AND COMPLEXITY REDUCTION FOR PROPOSED FAST INTRA PREDICTION ON THE HM10 MAIN PROFILE USING HEVC COMMON TEST CONDITION [29] WITH QPS AT 22, 27, 32, AND 37 FOR CLASS F SEQUENCES

Class & Sequences		Luma-Y	Chroma-U	Chroma-V	ΔT
F	BasketballDrillText	1.4%	2.3%	2.4%	55%
	ChinaSpeed	2.1%	2.3%	3.4%	60%
	SlideEditing	2.1%	1.6%	1.8%	57%
	SlideShow	1.7%	1.5%	2.0%	74%
Ave.		1.8%	1.9%	2.4%	61%

proposed fast intra mode decision scheme. On average, our proposed solution achieves 60% encoding time reduction for all intracoding with 1.0% BD-rate increase. In addition, we also perform the simulation using the optional Class F screen content sequences, where our proposed fast algorithm provides 61% encoding time reduction with 1.8% BD-rate increase, as shown in Table V.

Moreover, additional tests are carried out to study whether the proposed algorithm is applicable to other bit-rate ranges using different QPs rather than the QPs defined in the HEVC common test conditions [29]. Two typical bit-rate ranges are considered in this paper, i.e., high bit-rate range (noted as Prop.-HBR) using the QPs at 12, 17, 22, and 27, and low bit-rate range (noted as Prop.-LBR) using the QPs at 32, 37, 42, and 47. Other encoding parameters are kept the same as the common test condition. Simulation results are categorized in Table VII. For Prop.-HBR, our proposed algorithm provides averaged 55% encoding time reduction with just 0.8% BD-rate increase, while 68% encoding time reduction with 1.5% BD-rate increase for Prop.-LBR, both against the default intra mode decision of the HEVC reference implementation. Results further reveal that our proposed method works for different bit-rate ranges quite well.

In addition, we have also performed the experiments for other encoder settings, such as RA, LD with B picture (LD-B), and low-delay with P picture (LD-P) to showcase the efficiency of our proposed solution. On average, it shows 17%, 15%, and 15% complexity reduction with 1.1%, 0.8%, and 0.9% BD-rate loss for respective RA, LD-B, and LD-P configurations. As detailed in the HEVC common test condition [29], intraframes are placed at approximately 1-s intervals for RA setting, but only the first picture is encoded using intraframe for both LD-B and LD-P cases. For simplicity, only averaged numbers are listed for different classes of sequences, as summarized in Table VI.

TABLE VI

CODING EFFICIENCY AND COMPLEXITY REDUCTION FOR PROPOSED FAST INTRA PREDICTION ON THE HM10 MAIN PROFILE USING HEVC COMMON TEST CONDITION [29] WITH QPS AT 22, 27, 32, AND 37 FOR RA, LD-B, AND LD-P CONFIGURATIONS

		Luma-Y	Chroma-U	Chroma-V	ΔT
RA	Class A	1.0%	1.2%	0.6%	18%
	Class B	0.8%	1.2%	1.6%	16%
	Class C	1.3%	1.8%	1.8%	16%
	Class D	1.2%	1.4%	1.7%	15%
Ave.		1.1%	1.4%	1.4%	17%
LD-B	Class B	0.7%	0.8%	0.6%	15%
	Class C	1.0%	1.2%	1.4%	13%
	Class D	0.7%	0.7%	1.2%	12%
	Class E	1.3%	1.8%	1.2%	18%
Ave.		0.8%	1.1%	1.4%	15%
LD-P	Class B	0.7%	0.6%	0.8%	15%
	Class C	1.0%	1.0%	1.4%	14%
	Class D	0.8%	0.5%	0.8%	13%
	Class E	1.1%	2.1%	1.3%	19%
Ave.		0.9%	1.0%	1.0%	15%

B. Performance Comparison

As introduced in Section II, there are noticeable amount of works on fast intra mode decision proposed for HEVC. We have implemented several algorithms from the recent published works [17], [31]–[33]. We have found the results claimed by the these works are almost similar at the latest HM platform according to our experiments. It might be due to the fact that HEVC intra prediction does not involve significant changes after the publication of these papers. Simulations are performed using the common test conditions provided in [29], and the results are detailed in Table VII for the side-by-side comparison with our proposed algorithm. On average, Jiang's work [17] gives the least BD-rate increase at about 0.4% but only provides $\approx$ 20% encoding time reduction; Zhang's work [31] yields the similar encoder time reduction as our solution, but with significant performance loss, i.e., $\approx$ 5% BD-rate increase. As we can see, our proposed fast intra mode decision method (i.e., noted as Prop.-Overall) gives the state-of-the-art performance in terms of the tradeoff between the coding efficiency and the encoding complexity.

We additionally list the results in Table VII with only pRMS (noted as Prop.-pRMS), micro-level algorithms (noted as Prop.-Micro which is the combination of pRMS in Section III-B1 and early RDOQ skip in Section III-B2), and combined micro-level and macro-level algorithms

TABLE VII

PERFORMANCE COMPARISON BETWEEN THE PROPOSED SOLUTION AND THE EXISTING ALGORITHMS (WITH RESPECT TO THE LUMA BD-RATE LOSS ΔR AND ENCODING TIME REDUCTION ΔT) FOR SEQUENCE CLASSES A−E (WITH AVERAGED PERFORMANCE FOR ALL SEQUENCE CLASSES SHOWN IN LAST ROW)

	[17]		[31]		[32]		[33]		Prop.-pRMS		Prop.-Micro		Prop.-Overall		Prop.-LBR		Prop.-HBR	
	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT	ΔR	ΔT
A	0.6%	20%	5.9%	53%	1.0%	22%	0.6%	15%	0.3%	26%	0.6%	46%	0.8%	61%	1.4%	68%	0.6%	57%
B	0.7%	25%	6.7%	56%	0.8%	21%	0.5%	15%	0.4%	27%	0.8%	45%	0.9%	61%	1.5%	69%	0.6%	57%
C	0.7%	19%	3.5%	56%	1.5%	25%	0.5%	14%	0.5%	26%	0.9%	45%	1.1%	56%	1.5%	64%	0.9%	52%
D	0.9%	20%	4.4%	63%	1.7%	24%	0.8%	15%	0.7%	27%	0.8%	44%	1.0%	55%	1.4%	61%	1.0%	51%
E	0.9%	19%	-	-	1.3%	23%	0.9%	15%	0.4%	26%	1.1%	45%	1.5%	64%	2.0%	71%	1.0%	57%
	0.8%	21%	4.9%	57%	1.2%	23%	0.6%	15%	0.4%	26%	0.8%	45%	1.0%	60%	1.5%	68%	0.8%	55%

(Prop.-Overall) to show the performance aggregation by combing different tools proposed in our solution.

V. CONCLUSION

We propose a novel fast intra mode decision for HEVC intraencoding. The overall algorithm consists of two main aspects: one is the micro-level scheme where we develop the progressive rough mode search based on the Hadamard cost to selectively evaluate potential prediction modes (rather traversing all available modes as the default rough mode decision in the HEVC reference implementation), and the early RDOQ skip to further reduce the unlikely modes in the effective mode set; the other is the macro-level method where we early terminate the coding unit split if the estimated R-D cost of the all sub-CUs (through the available Hadamard costs and aggregated R-D costs of partial sub-CUs) is larger than the CU R-D cost (without split into sub-CUs).

Extensive simulations report the 60% encoding time reduction with 1.0% BD-rate increase using the HEVC common test condition over a cluster of different sequences compressed with different QPs. More results are also demonstrated at respective high bit-rate (with QPs at 12, 17, 22, 27) and low bit-rate (with QPs at 32, 37, 42, 47) scenarios, where complexity reduction is 55% and 68%; and BD-rate increase is 0.8% and 1.5%, respectively. In addition, experiments on other encoder settings, such as RA, LD-B, and LD-P, are carried out following the HEVC common test condition, with respective averaged 17%, 15%, and 15% encoder complexity reduction at $\approx$ 1% BD-rate increase. Meanwhile, compared with the other works, our solution demonstrates the state-of-the-art performance in terms of the tradeoff between BD-rate loss and complexity reduction (see Table VII).

All tests discussed so far are compliant with the HEVC main profile focusing on the 8-bit consumer video. Extra experiments are also performed for HEVC main10 profile focusing on the 10-bit professional video. Simulation results demonstrate almost the same encoder complexity reduction (slightly ±1% difference such as 60% to 61% or 59%) and BD-rate loss (slightly ±0.1% difference such as 1.0% to 0.9% or 1.1%) for all intra, random access and low-delay configurations, in comparison to the HEVC main profile compliant tests. These results are not included in this paper due to the space limitation, but are placed on a publicly accessible server at http://vision.poly.edu/∼zma03/eData/HEVCfastIntra.xls.

Class F sequences, also called screen content or non-camera captured videos, are included into the tests in this paper. Even though the complexity reduction is retained, we see that the BD-rate increase is larger than the conventional Class A to Class E camera captured videos, i.e., from averaged 1.0% to 1.8% for all intracoding (cp. Table IV and V). This is probably due to the quite different content distribution between videos in Class F and other classes. It requires substantial effort to fully understand the Class F videos and come up with a refined algorithm with a better complexity and coding efficiency tradeoff. This could be part of our next step research exploration. Such kind of work is very meaningful since ISO/IEC Moving Picture Expert Group already issued a draft call for proposal with the focus on the high efficiency screen content coding under the current HEVC framework [34].

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments to improve this manuscript significantly.

REFERENCES

[1] B. Bross, W.-J. Han, J.-R. Ohm, G.-J. Sullivan, Y.-K. Wang, and T. Wiegand, *High Efficiency Video Coding (HEVC) Text Specification Draft 10*, document JCTVC-L1003 (rev. 37), Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Mar. 2013.

[2] G.-J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[3] J.-R. Ohm, G.-J. Sullivan, H. Schwarz, T.-K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards: Including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.

[4] I.-K. Kim, J. Min, T. Lee, W.-J. Han, and J. Park, "Block partitioning structure in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1697–1706, Dec. 2012.

[5] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2012.

[6] *HEVC Reference Model* [Online]. Available: http://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/

[7] I. Kim, K. McCann, K. Sugimoto, B. Bross, and W.-J. Han, *High Efficiency Video Coding (HEVC) Test Model 10 (HM10) Encoder Description*, JCT-VC document L1002, Jan. 2013.

[8] Y. Piao, J. Min, and J. Chen, *Encoder Improvement of Unified Intra Prediction*, document JCTVC-C207, Oct. 2010.

[9] F. Pan, X. Lin, S. Rahardja, K. Lim, Z. Li, D. Wu, *et al.*, "Fast mode decision algorithm for intra prediction in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 813–822, Jul. 2005.

[10] A.-C. Tsai, J.-F. Wang, J.-F. Yang, and W.-G. Lin, "Effective subblock-based and pixel-based fast direction detections for H.264 intra prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 7, pp. 975–982, Jul. 2008.

[11] Y.-H. Huang, T.-S. Ou, and H. Chen, "Fast decision of block size, prediction mode, and intra block for H.264 intra prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 8, pp. 1122–1132, Aug. 2010.

[12] K. Bharanitharan, B.-D. Liu, J.-F. Yang, and W.-C. Tsai, "A low complexity detection of discrete cross differences for fast H.264/AVC intra prediction," *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1250–1260, Nov. 2008.

[13] C.-H. Tseng, H.-M. Wang, and J.-F. Yang, "Enhanced intra-4x4 mode decision for H.264/AVC coders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 8, pp. 1127–1132, Aug. 2006.

[14] A.-C. Tsai, A. Paul, J.-C. Wang, and J.-F. Wang, "Intensity gradient technique for efficient intra-prediction in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, pp. 694–698, May 2008.

[15] H. Zeng, K.-K. Ma, and C. Cai, "Hierarchical intra mode decision for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 907–912, Jun. 2010.

[16] K. Choi and E. Jang, "Fast coding unit decision method based on coding tree pruning for high efficiency video coding," *Opt. Eng.*, vol. 51, no. 3, p. 030502, Mar. 2012.

[17] W. Jiang, H. Ma, and Y. Chen, "Gradient based fast mode decision algorithm for intra prediction in HEVC," in *Proc. CECNet*, Apr. 2012, pp. 1836–1840.

[18] X. Shen, L. Yu, and J. Chen, "Fast coding unit size selection for HEVC based on Bayesian decision rule," in *Proc. PCS*, May 2012, pp. 453–456.

[19] G. Tian and S. Goto, "Content based hierarchical fast coding unit decision algorithm for HEVC," in *Proc. PCS*, May 2012, pp. 56–59.

[20] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective CU size decision method for HEVC encoders," *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.

[21] H. Zhang and Z. Ma, "Early termination schemes for fast intra prediction in high-efficiency video coding," in *Proc. IEEE ISCAS*, May 2013, pp. 45–48.

[22] L. Zhao, L. Zhang, S. Ma, and D. Zhao, "Fast mode decision algorithm for intra prediction in HEVC," in *Proc. IEEE VCIP*, Nov. 2011, pp. 1–4.

[23] H. Zhang and Z. Ma, "Fast intra prediction for high efficiency video coding," in *Proc. PCM*, Dec. 2012, pp. 568–577.

[24] Y.-H. Tan, C.-H. Yeo, H.-L. Tan, and Z.-G. Li, "On residual quad-tree coding in HEVC," in *Proc. IEEE MMSP*, Oct. 2011, pp. 1–4.

[25] S.-W. Teng, H.-M. Hang, and Y.-F. Chen, "Fast mode decision algorithm for residual quadtree coding in HEVC," in *Proc. IEEE VCIP*, Nov. 2011, pp. 1–4.

[26] Y. Zhang and M. Zhao, "An adaptive RQT mode selection algorithm for HEVC," in *Proc. 5th IEEE Int. Congr. Image Signal Process.*, Oct. 2012, pp. 173–177.

[27] R. Li, B. Zeng, and M. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 4, pp. 438–442, Aug. 1994.

[28] S. Zhu and K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 287–290, Feb. 2000.

[29] F. Bossen, *Common HM Test Conditions and Software Reference Configurations*, JCT-VC document L1100, Jan. 2013.

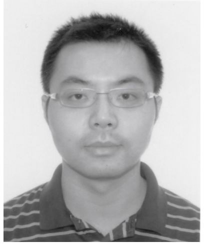
[30] G. Bjontegaard, *Calculation of Average PSNR Differences Between R-D Curves*, document VCEG-M33, ITU-T VCEG 13th Meeting, Apr. 2001.

[31] Y. Zhang, Z. Li, and B. Li, "Gradient-based fast decision for intra prediction in HEVC," in *Proc. IEEE VCIP*, Nov. 2012, pp. 1–6.

[32] S. Yan, L. Hong, W. He, and Q. Wang, "Group-based fast mode decision algorithm for intra prediction in HEVC," in *Proc. Int. Conf. Signal Image Technol. Internet Based Syst.*, 2012, pp. 225–229.

[33] M. Zhang, C. Zhao, and J. Xu, "An adaptive fast intra mode decision in HEVC," in *Proc. IEEE VCIP*, Nov. 2012, pp. 221–224.

[34] H. Yu, K. McCann, R. Cohen, and P. Amon, *Draft Call for Proposals for Coding of Screen Content and Medical Visual Content*, ISO/IEC JTC1/SC29/WG11 N13829, Jul. 2013.



Hao Zhang (M'06) received the Ph.D. degree in electrical and computer engineering from Polytechnic University (now Polytechnic Institute of New York University), New York, NY, USA, in 2006.

He is an Associate Professor with the School of Information Science and Technology, Central South University, Changsha, China. His research interests include video coding and image processing.

Zhan Ma (S'06–M'11) received the B.S. and M.S. degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2004 and 2006, respectively, and the Ph.D. degree in electrical engineering from Polytechnic University (now Polytechnic Institute of New York University), New York, NY, USA, in 2011.

He was an Intern with the Thomson Corporate Research Laboratory, Princeton, NJ, USA; Texas Instruments, Dallas, TX, USA; and Sharp Laboratories of America, Camas, WA, USA, in 2008, 2009, and 2010, respectively. Since 2011, he has been with Samsung Research America at Dallas, Richardson, TX, USA, as a Senior Standards Researcher. His research interests include next-generation video coding standardization, video fingerprinting, and video signal modeling.