Fast Mode Decision Technique for HEVC Intra Prediction Based on Reliability Metric for Motion Vectors

Chihiro TSUTAKE¹, Nonmember, Yutaka NAKANO¹, and Toshiyuki YOSHIDA¹, Members

SUMMARY
This paper proposes a fast mode decision technique for intra prediction of High Efficiency Video Coding (HEVC) based on a reliability metric for motion vectors (RMMV). Since such a decision problem can be regarded as a kind of pattern classification, an efficient classifier is required for the reduction of computation complexity. This paper employs the RMMV as a classifier because the RMMV can efficiently categorize image blocks into flat (uniform), active, and edge blocks, and can estimate the direction of an edge block as well. A local search for angular modes is introduced to further speed up the decision process. An experiment shows the advantage of our technique over other techniques.

key words: H.265/HEVC, intra prediction, mode decision, reliability metric, coding efficiency, computational complexity

1. Introduction

Current video coding standards employ a hybrid coding framework, in which spatio-temporal redundancies are removed by a combination of prediction and transformation techniques [1], [2]. The H.265/HEVC (High Efficiency Video Coding) and its predecessor H.264/AVC (Advanced Video Coding) standard utilize intra prediction techniques, where each pixel in an intra-coding block is predicted from its neighboring and already reconstructed reference pixels [3]–[5].

As briefly reviewed in the next section, the intra prediction of HEVC offers a total of 35 prediction modes comprising planar, DC, and 33 angular modes. In addition to increasing the number of modes in AVC, HEVC allows recursive splitting of Coding Units (CUs), which generates Prediction Units (PUs) of sizes ranging from 64 \times 64 to 4 \times 4 in intra coding. The optimal determination of the prediction mode and the PU size basically requires an exhaustive search of possible combinations of the mode and the size. This is why a fast mode and/or PU-size decision techniques are required in an actual HEVC encoder.

The HEVC Test Model (HM), which is the reference software of HEVC, determines the optimal prediction mode based on two phases, Rough Mode Decision (RMD) and Rate-Distortion Optimization (RDO). Since a small number of candidates out of 35 modes are first selected by the RMD before full rate-distortion optimization in the RDO, HM can lower computation costs compared with the exhaustive search of 35 modes in the time-consuming RDO process.

Recently, much effort has been made for further easing the computation complexity of HM intra coding, which is roughly classified into two categories, fast block partitioning techniques [6]–[9] and fast mode decision techniques [10]–[18]. This paper focuses on fast mode decision techniques, and proposes a novel technique that outperforms the existing approaches. As shown in the next section, many of the fast mode decision techniques are based on a combination of flat/edge block classification and edge direction estimation.

In Ref. [19], the authors proposed a reliability metric for motion vectors (RMMV) to evaluate the reliability of motion vectors by a block matching technique. As demonstrated in Sect. 3, the RMMV also can efficiently classify image blocks into flat, active, and edge blocks, and can estimate the edge direction. These properties of the RMMV motivate us to apply it to the fast mode decision problem. A local search for angular modes is also introduced to further reduce computation complexity.

The rest of this paper is organized as follows. Section 2 briefly reviews the intra coding framework of the HEVC/HM and related works. Section 3 explains the definition and properties of the RMMV by referring to some calculation examples of the RMMV. Then, Sect. 4 elaborates the proposed fast mode decision technique, and Sect. 5 gives some experimental results and comparisons to validate the efficiency of our technique. Finally, Sect. 6 concludes this paper.

2. HEVC Intra Prediction and Related Works

2.1 HEVC Intra Prediction [3]–[5]

HEVC employs a quadtree structure of coding blocks referred to as Coding Tree Units (CTUs), which are recursively divided into smaller blocks called CUs as shown in Fig. 1. Each leaf block of this tree structure is referred to as a PU, which is utilized as a unit of prediction. In HEVC intra coding, the size of a PU ranges from 64 \times 64 to 4 \times 4 depending on the texture/activity of each CU.

The intra coding of HEVC offers a total of 35 prediction modes comprising the planar, DC, and 33 angular modes (Fig. 2 (a)). Each mode generates a prediction block from its neighboring, already reconstructed reference pixels. The DC mode generates a uniform prediction block by using...
the average value of the reference pixels, whereas the planar mode makes a prediction based on a bilinear interpolation over the four corner reference pixels. The remaining modes, i.e., the angular modes, are applied to directional PUs, such as an edge block. As shown in Fig. 2 (a), the angular modes extrapolate a linearly interpolated version of two adjacent reference pixels along with the direction determined by the mode number.

While a single prediction mode out of 35 is selected for each intra PU, an actual prediction is carried out based on transform units (TUs), which are subblocks derived by the quadtree-based partitioning for the CUs. An $N \times N$ TU makes use of a total of $4N + 1$ reference pixels indicated by the dark circles in Fig. 2 (b), and yields prediction pixels according to the mode of the PU to which the TU belongs. However, all reference pixels are not always available; for example, in the cases that the target TU is located at frame/slice boundaries, or that the reference pixels have not yet been reconstructed. In these cases, the closest available reference pixel is substituted for unavailable ones [4], which makes the optimal determination of the prediction mode very complicated, as discussed again in Sect. 4.

Since RD based intra coding requires optimal determinations of not only the prediction mode but also the PU size, an exhaustive search for all possible combinations of the mode and the size is necessary. Such a search would incur an unrealistic computation cost. This is the reason why numerous of techniques have been proposed to reduce the computation complexity for HEVC intra coding.

2.2 Intra Coding Implementation in the HM [4]

The optimal intra prediction mode for a PU can be determined by minimizing the full RD cost function

$$ J_{\text{mode}} = \text{SSE} + \lambda_{\text{mode}} B_{\text{mode}}, $$

(1)

where SSE represents the distortion in terms of the sum of squared error between the original and the reconstructed pixels, $B_{\text{mode}}$ is the total bit cost for encoding the current single PU, and $\lambda_{\text{mode}}$ is a Lagrange multiplier.

Since an evaluation of Eq. (1) requires a very large cost, the HEVC test model HM employs the following two-phase algorithm for determining the intra prediction mode. In the first phase, referred to as the Rough Mode Decision (RMD), the cost function

$$ J_{\text{pred, satd}} = \text{SATD} + \lambda_{\text{pred}} B_{\text{pred}}, $$

(2)

is evaluated over every 35 prediction mode for the current PU to find the least $N$ modes as candidates. In Eq. (2), SATD is the sum of absolute difference between the Hadamard-transformed original and the predicted pixels, $B_{\text{pred}}$ is the bit cost for encoding the mode information of the current PU, and $\lambda_{\text{pred}}$ is a Lagrange multiplier. The number of candidates $N$ depends on the PU size, and is defined as 8 for $4 \times 4$ and $8 \times 8$ PUs and 3 for the other sizes, as shown in the column “HM” in Table 1. Then, Most Probable Modes (MPMs) are included in the candidate list to exploit the spatial correlation of the optimal prediction mode [10]. In the second phase called Rate-Distortion Optimization (RDO), $J_{\text{mode}}$ in Eq. (1) is evaluated for each candidate in Table 1 to determine the optimal prediction mode.

Because the calculation cost of Eq. (2) is much smaller than that of Eq. (1), this algorithm can reduce the computation complexity compared with an exhaustive search of 35 modes in the time-consuming RDO process. Recently, much effort has been made to further reduce the complexity
of HEVC intra coding, which is briefly reviewed in the next subsection.

2.3 Review of Related Works

Fast intra coding techniques for HEVC can be roughly classified into two categories: fast block partitioning techniques [6]–[9] and fast mode decision techniques [11]–[18] as discussed in Sect. 1. Since this paper focuses on the fast mode decision techniques, this section briefly reviews some typical approaches for these techniques.

The fast mode decision techniques basically reduce the number of candidate modes in the RMD and/or the RDO. Refs. [11]–[16] calculate the gradient direction of a PU to derive a direction along which the pixel values are almost uniform. In particular, Refs. [11]–[13] first calculate the pixel-wise gradient direction, which is converted into the corresponding angular mode, and then select a small set of candidates for the RMD based on a histogram analysis for the pixel-wise angular modes [11], [12], or a kernel density estimation approach [13]. The other gradient-based techniques [14]–[16] first estimate a rough gradient direction among 0°, 45°, 90°, and 135°, and then select some neighboring modes centered at the rough direction for the RMD. In contrast, Refs. [17], [18] do not estimate the gradient direction, but adaptively control the number of candidate modes in the RMD and/or the RDO phases.

It can be observed from this review that these techniques are based on the following fundamental concepts:

a) PU classification for suitably selecting the planar, DC, or angular modes
b) The gradient direction computation to estimate the optimal angular mode
c) Adaptive control for the number of candidate modes in the RMD and/or the RDO.

This paper proposes a novel fast mode decision technique that is also based on concepts a)–c). In our technique, a) and b) are realized by using the RMMV, which is briefly explained in the next section, while c) is carried out by introducing a local search for the angular mode with the least $J_{pred,SATD}$ in the RMD.

3. Reliability Metric for Motion Vectors (RMMV)

3.1 Derivation of the RMMV [19]

The authors have proposed a reliability metric for motion vectors (RMMV), which was originally derived for evaluating “how an estimated MV is reliable”. This section briefly reviews the RMMV by referring to a typical MV-based inter-frame video coding technique, and gives some calculation examples.

MVs are utilized in the hybrid video coding framework to represent translational motion between the reference frame $g(x, y)$ and the current frame $f(x, y)$ to be encoded, where $x$ and $y$ are integer indices. In motion estimation, the MV $(v_x, v_y)$ for a target block $B$ on $f$ can be obtained with the so-called Block Matching (BM) technique:

$$ (v_x, v_y) = \arg \min_{(v'_x, v'_y) \in \text{SR}} L(f(x, y), g(x - v'_x, y - v'_y)), $$  \hspace{1cm} (3)

where $L(f, g)$ represents an error criterion between $f$ and $g$ on block $B$, and SR denotes the search range for the MV. The mean squared error criterion

$$ L = \sum_{(x, y) \in B} (f(x, y) - g(x - v'_x, y - v'_y))^2 $$  \hspace{1cm} (4)

is utilized in this section.

Let us assume that block $B$ on the current frame $f$ is a translated and noise-corrupted version of a block on the reference frame $g$:

$$ f(x, y) = g(x - V_x, y - V_y) + n(x, y) \quad ((x, y) \in B), $$  \hspace{1cm} (5)

where $(V_x, V_y)$ represents the true translational motion, and $n(x, y)$ is Gaussian white noise. Due to the presence of the additive noise $n(x, y)$, $(v_x, v_y)$ which minimizes Eq. (3) does not coincide with the true motion $(V_x, V_y)$ in Eq. (5).

In Ref. [19], the authors defined the error $(\Delta v_x, \Delta v_y)$ between the true MV $(V_x, V_y)$ and the $(v_x, v_y)$ that minimizes Eq. (3) as

$$ \left\{ \begin{array}{l} \Delta v_x = V_x - v_x \\ \Delta v_y = V_y - v_y \end{array} \right., $$  \hspace{1cm} (6)

and then obtained the ensemble average of $\Delta^2_v$, $\Delta^2_n$, and $\Delta \Delta_g$ by using the fact that $n(x, y)$ is Gaussian white noise. Finally, the error in direction $\theta$ is defined as

$$ \Delta(\theta) = \Delta \cos \theta + \Delta \sin \theta, $$  \hspace{1cm} (7)

and the ensemble average of $\Delta(\theta)$, denoted as $R(\theta)$, is derived:

$$ R(\theta) = C \cdot \frac{\sin^2 \theta \sum_B f_x^2 + \cos^2 \theta \sum_B f_y^2 - 2 \sin \theta \cos \theta \sum_B f_x f_y}{\sum_B \sum_B f_x^2 \sum_B f_y^2 - \left( \sum_B f_x f_y \right)^2}, $$  \hspace{1cm} (8)

where

$$ f_x = \partial f / \partial x, \quad f_y = \partial f / \partial y, $$

and $C$ is a constant in terms of the noise power.

Since Eq. (8) represents the mean squared error between the true and estimated MVS under the presence of additive Gaussian white noise, the value of $R(\theta)$ is presumed to indicate the reliability of the MV in the direction $\theta$, that is, how accurately the MV can be estimated in the direction $\theta$. Note here that a larger value of $R(\theta)$ indicates that the estimated MV is less reliable in the direction $\theta$, as demonstrated in the next subsection.

3.2 Examples of the RMMV

Figure 3 (f) gives examples of $R(\theta)$ calculated for $16 \times 16$
blocks (a)–(e). In the rest of this paper, constant C in Eq. (8) is fixed to 1, and \( \theta_{\text{max}} \) (\( \theta_{\text{min}} \)) represents the direction that gives the largest (smallest) value \( R_{\text{max}} \) (\( R_{\text{min}} \)) of the curve \( R(\theta) \).

For a block as flat as Fig. 3 (a), the MV cannot be estimated precisely in any direction, which leads to a large estimation error in general. The curve \( R(\theta) \) for block (a) in Fig. 3 (f) illustrates this context, where the value of \( R(\theta) \) is large for any direction \( \theta \). On the other hand, for an active or highly textured block (d), the value of \( R(\theta) \) is small for any direction, indicating that the MV of the block can be obtained with a small error in an arbitrary direction. These examples suggest that flatness or uniformity of a block can be rated by the RMMV, e.g., by comparing \( R_{\text{min}} \) with a predetermined threshold value.

For edge or directional blocks, shown Figs. 3 (b) and (c), the MV component parallel to the edge direction cannot be estimated although the orthogonal component to the edge can be precisely obtained: this is known as the “aperture problem”. Such a characteristic in motion estimation is observed in Fig. 3 (f), where the value of \( R(\theta) \) is large (small) in the direction parallel (orthogonal) to the edge for blocks (b) and (c). It should be noted that \( \theta_{\text{max}} \) hence gives the edge direction for an edge block. Furthermore, the “edge intensity” of block (b) is higher than that of (c), which is also estimated from Fig. 3 (f) as the variation of the curve \( R(\theta) \); the variation of the “stronger” edge (b) is larger than that of edge (c). Since the variation of \( R(\theta) \) can be evaluated by

\[
r = \frac{R_{\text{max}}}{R_{\text{min}}},
\]

the “edge intensity” can be assessed by the value \( r \) of a target block.

### 3.3 Normalization of the RMMV

Since, in general, MVs can be estimated with higher accuracy for larger blocks of a uniform texture, the RMMV in Eq. (8) tends to give a smaller value for a larger block. Such a property is demonstrated by blocks (c) and (e) in Fig. 3, where block (e) comprises a texture similar to (c) but is twice the size of (c). The dependency of \( R(\theta) \) on the block size is observed from RMMV curves (c) and (e) in Fig. 3 (f).

In this paper, since our task is to evaluate the uniformity or directionality of a block on a single scale irrespective of the block size, the normalized version of the RMMV in terms of the block size

\[
R(\theta) = N_xN_y \cdot R(\theta)
\]

is utilized, where \( N_x \) and \( N_y \) are the width and height of block \( B \), respectively.

### 3.4 Summary and Comments

According to the discussion in Sects. 3.1 and 3.2, the properties of the RMMV can be summarized as follows: for a target block,

(1) “flatness” can be evaluated by the value \( R_{\text{min}} \)
(2) “edge intensity” can be evaluated by the ratio \( r \).

These properties suggest that the RMMV can be an efficient block classifier for a fast intra mode decision technique, although it was originally derived for motion estimation in inter coding. The proposed technique makes use of the RMMV to reduce the calculation cost.

Before concluding this subsection, a few comments are made for the calculation of \( R(\theta) \), \( R_{\text{max}} \), and \( R_{\text{min}} \). Equation (8) requires the partial derivatives \( f_x \) and \( f_y \) for the target block \( f \). The simplest way to calculate the partial derivatives on a discrete image is to replace them with the finite differences in terms of \( x \) and \( y \) as shown in Fig. 4 (a). Unfortunately, such simple finite differences yield mutually different reference points for \( x \) and \( y \) (depicted as black circles). These different reference points lead to a large calculation error, in particular, for the directions \( \theta_{\text{max}} \) and \( \theta_{\text{min}} \). To avoid such an error, Ref. [19] employs the masks in Fig. 4 (b) and (c) for \( f_x \) and \( f_y \), respectively. Note that, in this case, the reference points for \( f_x \) and \( f_y \) coincide with each other.

Ref. [19] gives another expression for the RMMV in Eq. (8), from which \( R_{\text{max}} \), \( R_{\text{min}} \), \( \theta_{\text{max}} \), and \( \theta_{\text{min}} \) can be readily calculated. By taking the normalization in Eq. (10) into account, these are written as

\[
R_{\text{max}} = \frac{CN_xN_y}{a - ||p||} \left( \theta_{\text{max}} = \frac{1}{2} \arg p + \frac{\pi}{2} \right)
\]

\[
R_{\text{min}} = \frac{CN_xN_y}{a + ||p||} \left( \theta_{\text{min}} = \frac{1}{2} \arg p \right),
\]
Fig. 4 Calculation of the partial derivatives $f_x$ and $f_y$.

(a) $f_x$ and $f_y$ leading to mutually different reference points

(b) $f_x$

(c) $f_y$

Fig. 5 Partitioning of the $r$-$R_{\text{min}}$ plane.

\( a = \sum_B f_x^2 + f_y^2 \)

\( p = \left( \sum_B f_x^2 - f_y^2, \sum_B f_x f_y \right) \)  \( \ldots \) (13)

4. Proposed Technique

Based on the preparations thus far, this section elaborates the proposed technique for fast intra mode decision in the HEVC. Our technique exploits all of the ideas a)–c) in Sect. 2.3: the RMMV is utilized to realize a) and b), while c) is carried out by a local search of angular modes.

In the proposed technique, $R_{\text{min}}$, $\theta_{\text{max}}$, and $r$ are first calculated for each PU, and then the PU is mapped onto the corresponding point on the $r$-$R_{\text{min}}$ plane. As shown in Fig. 5, if the $r$-$R_{\text{min}}$ plane is partitioned into regions 1–3 with the threshold values $Th_r$ and $Th_{R_{\text{min}}}$, the properties of the RMMV guarantee that regions 1–3 basically include directional, flat, and active PUs, respectively. The edge direction can be obtained from the $\theta_{\text{max}}$ for a PU in region 1, which is then associated with the closest angular mode denoted by $A$. Figure 6 gives examples of the $(r, R_{\text{min}})$ distribution together with a typical PU in each region, from which we can confirm that the RMMV can efficiently classify PUs into one of the regions 1, 2 and 3.

The simplest way to exploit such a block classification is to reduce the number of candidate modes for the RMD/RDO individually in each region. For example, for a PU in region 1, the angular mode $A$ can be selected alone as the optimal mode, and the RMD and RDO phases are skipped. In the same way for a PU in region 2, the planar and DC modes are selected and sent to the RDO.

Such a simple technique, however, does not work well because all the necessary reference pixels are not always available for each TU. As discussed in Sect. 2.1, missing reference pixels are replaced with the nearest available ones. Therefore, this process greatly reduces the correlation between pixels to be predicted and those to be referred, result-

\[\text{As pointed out in Ref. [17], since an optimal mode estimation for the planar and DC modes is actually difficult, both modes should initially be selected together as candidates in the RMD/RDO.}\]
ing in poor prediction. In such a case, the angular mode $A$ is not necessarily optimal in the directional PUs, and a planar, DC, or another angular mode is possibly selected as the best mode. This fact suggests that a winner-takes-all decision rule does not always work well, and that a set of candidate modes is still required for the RDO. To cope with such a requirement while keeping the calculation cost as low as possible, a local search of angular modes has been introduced in the proposed technique.

Based on the observation and discussion so far, this paper proposes the following fast mode decision technique using the RMMV. For each PU, the following process takes place:

1. **Initialization:** $R_{\text{min}}$ and $r$ are calculated to determine the region to which the PU belongs, and the closest angular mode $A$ is identified from the edge direction $\theta_{\text{max}}$. The planar, DC, and the angular mode $A$ are selected as candidates for the RMD.

2. **RMD:** To find the $N$ best candidates for the RDO, $J_{\text{pred SATD}}$ is first evaluated for each candidate in step 1, and then starting with these values, a local search for angular modes is initiated if necessary. The number of RDO candidates $N$ is dependent on the PU size and the regions in Fig. 5, as listed in Table 1 (See below).

3. **MPM:** Just as in the HM, the most probable modes (MPMs) are included in the RDO candidate list.

4. **RDO:** The optimal mode is determined by an evaluation of $J_{\text{mode}}$ for each of the RDO candidates.

In step 2, the number of RDO candidates $N$ in each

| PU size | HM | Proposed | Ref.[16] |
|---------|----|----------|----------|
| 64 × 64 | 3  | 1 1 1     | 1 1 1     |
| 32 × 32 | 3  | 1 1 1     | 1 1 1     |
| 16 × 16 | 3  | 1 1 1     | 1 1 1     |
| 8 × 8  | 8  | 1 1 3     | 3 3 3     |
| 4 × 4  | 8  | 3 1 1     | 1 1 3     |

**Fig. 7** Detailed algorithm in step 2.
region was determined on the following basis. For flat or direction PUs in regions 1 and 2, \( N = 1 \) is sufficient for all sizes except for 4 \( \times 4 \) PUs in region 1. Since 4 \( \times 4 \) PU is not large enough to strictly determine its direction, additional modes should be included in the RDO candidate list. \( N = 3 \) was thus selected for 4 \( \times 4 \) PUs in region 1. On the other hand, \( N \) was reduced to 1 for the same size of PU in region 3. This is because the best and the second best modes in the RMD give almost similar prediction performances for 4 \( \times 4 \) PUs in region 3, which suggests that \( N = 1 \) is sufficient for that size in this region.

Figure 7 shows a detailed algorithm of the local search in step 2, which selects the \( N \) best candidates among the planar, DC, and the local-minimum angular modes. In (1) and (2) in Fig. 7, \( J_{\text{pred, SATD}} \) is evaluated for each candidate selected in step 1, i.e., for the planar and DC modes and angular mode, and then they are sorted in ascending order. Let \( m_1 \), \( m_2 \), and \( m_3 \) denote the modes in the sorted order. If the planar and DC modes outperform mode \( A \) in (3), the two modes \( m_1 \) and \( m_2 \) are put into the RDO candidate list in (10). Note that, in this case, the local search is skipped. Otherwise, the local search is started from (4). If \( J_A \) is a local minimum in (5), \( m_1 \) and \( m_2 \) are included in the list in (10). If \( J_A \) is not the case, the local search is continued in (6)–(8) to find the local minimum mode \( A_m \) in \( A \pm k (k = 1, 2, 3, 4) \). Note that in (7) and (8), \( A + 4 (A - 4) \) is determined as the local minimum \( A_m \) if \( J_{A+4} \leq J_{A+3} \leq J_{A+2} \leq J_{A+1} \) and \( J_{A-4} \leq J_{A-3} \leq J_{A-2} \leq J_{A-1} \). Finally, the \( N \) best modes are selected as the RDO candidates in (9) and (10).

5. Experimental Results

To verify its effectiveness, the proposed technique was implemented in the HM 14.1 encoder, and several test sequences were encoded in the intra-only Main profile.

With the original HM 14.1 encoder as an anchor method, the computation complexity and coding efficiency were evaluated, respectively in terms of the reduction ratio of the encoding time

\[
\Delta T = \frac{T_{\text{HM14.1}} - T_{\text{target}}}{T_{\text{HM14.1}}} \quad [%] \quad (14)
\]

and the BD-bitrate (BDBR) [%] (Luma-Y) with quantization parameters 22, 27, 32, and 37. In all the coding experiments described in this section, the CTU was fixed to 64\( \times \)64 with a maximum depth level of 4, and all the sequences were encoded every 10 frames. The total computation time \( T_{\text{target}} \) in Eq. (14) includes the time for calculating the RMMV in each PU, which is actually negligible.

The threshold values \( T_h \) and \( T_{\text{Rmin}} \) are first determined empirically in the next subsection.

5.1 Determination of the Threshold Values

The test sequences listed in Table 2 were encoded by the proposed technique with \( T_h \) and \( T_{\text{Rmin}} \) varied from 10 to 40 and 0.02 to \( \infty \), respectively, with a suitable step size. BDBR and \( \Delta T \) were averaged over all the sequences for each combination of \( T_h \) and \( T_{\text{Rmin}} \), as shown in Fig. 8.

![Fig. 8 Averaged BDBR and \( \Delta T \) with varied \( T_h \) and \( T_{\text{Rmin}} \). The number accompanying each curve represents the value of \( T_{\text{Rmin}} \). The large circle is the selected indicates the selected point \( T_h = 20.0 \) and \( T_{\text{Rmin}} = 0.1 \).](image)

| Test Seq.  | Size    | Proposed | Ref. [15] (our impl.) | Ref. [15] | Ref. [16] (our impl.) | Ref. [16] |
|-----------|---------|----------|-----------------------|----------|----------------------|----------|
| Traffic   | 2560\( \times \)1600 | 1.27     | 40.98                 | 2.29     | 32.98                | 1.69     |
| PeopleOnStreet | 2560\( \times \)1600 | 1.41     | 41.26                 | 2.21     | 31.94                | 1.5     |
| Kimonol   | 1920\( \times \)1080 | 1.03     | 44.07                 | 1.52     | 33.82                | 1.23     |
| ParkScene | 1920\( \times \)1080 | 0.95     | 42.24                 | 1.07     | 33.03                | 1.03     |
| Cactus    | 1920\( \times \)1080 | 1.99     | 42.02                 | 2.46     | 32.33                | 2.19     |
| BQTerrace | 1920\( \times \)1080 | 1.16     | 41.76                 | 1.49     | 33.93                | 1.77     |
| BasketballDrive | 1920\( \times \)1080 | 1.83     | 43.02                 | 2.96     | 34.06                | 2.22     |
| RaceHorses| 832\( \times \)480  | 1.55     | 41.56                 | 2.35     | 31.01                | 1.60     |
| PartyScene| 832\( \times \)480  | 2.26     | 39.52                 | 2.08     | 31.93                | 2.13     |
| BasketballDrill| 832\( \times \)480  | 2.36     | 39.28                 | 5.07     | 29.62                | 5.13     |
| RaceHorses| 416\( \times \)240  | 2.94     | 39.42                 | 3.32     | 31.71                | 2.90     |
| BQSquare  | 416\( \times \)240  | 3.84     | 40.56                 | 3.39     | 31.87                | 3.84     |
| FourPeople| 1280\( \times \)720 | 1.59     | 41.73                 | 2.59     | 33.82                | 2.16     |
| Johnny    | 1280\( \times \)720 | 2.23     | 43.64                 | 3.25     | 33.54                | 3.15     |
| KristenAndSara | 1280\( \times \)720 | 1.89     | 42.59                 | 3.29     | 33.61                | 2.99     |

"—" indicates that the value was not given in the corresponding reference.
angular modes was added to the rough mode decision phase in our technique. The local search for each PU size by our technique and Ref. [16]. Table 2 lists the results.

As shown in Table 1, our technique requires a lower number of J_{mode} evaluations than does Ref. [16]. This is one of the reasons that our technique achieves higher ΔT. Table 3 lists the average number of J_{pred,SATD} evaluations for each PU size by our technique and Ref. [16]. These values give another reason that our technique further reduces the computational complexity than does Ref. [16].

The efficiency improvement with our technique can be attributed to the fact that our technique can estimate a suitable prediction mode in particular for PUs in regions 1 and 2, which means that our technique has little impact on region 3. However, the improvements in regions 1 and 2 are so large that the total efficiency can be also improved in total, as can be seen in Table 1.

### 6. Conclusions

This paper has proposed a novel fast intra mode decision technique for HEVC. The reliability metric for motion vectors has been employed as a block classifier and an edge direction estimator in our technique. The local search for angular modes was added to the rough mode decision phase to reduce the number of evaluations of the SATD-based cost function. By applying these ideas, our technique achieves better coding efficiency while lowering the computation complexity over the previously proposed techniques.

Although this paper focused on fast mode decision techniques alone, an application of our technique together with a fast block partitioning approach can further reduce the computation complexity as in Ref.[20]–[23], which is left for future work.

### References

[1] Y. Sakai and T. Yoshida, Image Information Coding, Ohmsha, Tokyo, 2001 (in Japanese).
[2] T. Yoshida, T. Suzuki, and T. Hiroaki, Image Coding, Corona Publishing, Tokyo, 2008 (in Japanese).
[3] 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.
[4] T. Suzuki, S. Takamura, and T. Chujoh ed., H.265/HEVC Textbook, Impress Japan, Tokyo, 2013 (in Japanese).
[5] J. Lainema, F. Bossen, W.-J. Han, J. Min, K. Ugur, “Intra coding of the HEVC standard,” IEEE Trans. Circuits Syst. Video Technol., vol.22, no.12, pp.1792–1801, Dec. 2012.
[6] S. Cho and M. Kim, “Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding,” IEEE Trans. Circuits Syst. Video Technol., vol.23, no.9, pp.1555–1564, Sept. 2013.
[7] J. Kim, Y. Choe, and Y.-G. Kim, “Fast coding unit size decision algorithm for intra coding in HEVC,” Proc. IEEE Int’l Conf. on Consumer Electron., Las Vegas, USA, pp.637–638, Jan. 2013.
[8] M.U. K. Khan, M. Shafique, and J. Henkel, “An adaptive complexity reduction scheme with fast prediction unit decision for HEVC intra encoding,” Proc. IEEE Int’l Conf. on Image Processing, Melbourne, Australia, pp.1578–1582, Sept. 2013.
[9] L. Shen, Z. Zhang, and Z. Liu, “Effective CU size decision for HEVC intracoding,” IEEE Trans. Image Process., vol.23, no.10, pp.4232–4241, Oct. 2014.
[10] L. Zhao, L. Zhang, S. Ma, and D. Zhao, “Fast mode decision algorithm for intra prediction in HEVC,” Proc. IEEE Visual Commun. and Image Processing, Tainan, Taiwan, pp.1–4, Nov. 2011.
[11] W. Jiang, H. Ma, and Y. Chen, “Gradient based fast mode decision algorithm for intra prediction in HEVC,” Proc. 2nd Int’l Conf. on Consumer Electron., Communications and Networks, Yichang, China, pp.1836–1840, April 2012.
[12] A. BenHajyoussef, T. Ezzedine, and A. Bouallégue, “Fast gradient based intra mode decision for high efficiency video coding,” Int’l J. Emerging Trends & Technology in Computer Science, vol.3, no.3, pp.223–228, May-June 2014.
[13] G. Chen, Z. Liu, T. Ikenaga, and D. Wang, “Fast HEVC intra mode decision using matching edge detector and kernel density estimation alike histogram generation,” Proc. IEEE Int’l Symp. on Circuits and Syst. Beijing, China, pp.53–56, May 2013.
[14] C.-M. Fang, Y.-T. Chang, W.-H. Chung, “Fast intra mode decision for HEVC based on direction energy distribution,” Proc. IEEE 17th Int’l Symp. on Consumer Electron., pp.61–62, Hsinchu, Taiwan, June 2013.
[15] T.L. da Silva, L.V. Agostini, and L.A. da Silva Cruz, “HEVC intra prediction acceleration based on texture direction and prediction unit modes reuse,” APSIPA Trans. Signal and Information Processing, vol.3, e13, pp.1–13, Dec. 2014.
[16] Y. Yao, X. Li, and Y. Lu, “Fast intra mode decision algorithm for HEVC based on dominant edge assent distribution,” Multimedia Tools and Applications, published online 29 Nov. 2014, Nov. 2014.
[17] M. Zhang, C. Zhao, and J. Xu, “An adaptive fast intra mode decision...
in HEVC,” Proc. IEEE Int’l Conf. Image Processing, Orland, USA, pp.221–224, Sept. 2012.

[18] Y. Shi, O.C. Au, H. Zhang, X. Zhang L. Jia, W. Dai, and W. Zhu, “Local saliency detection based fast mode decision for HEVC intra coding,” Proc. IEEE 15th Int’l Workshop on Multimedia Signal Processing, Pula, Croatia, pp.429–433, Sept. 2013.

[19] T. Yoshida, A. Miyamoto, and Y. Sakai, “A reliability metric for motion vectors in moving pictures and its application,” IEICE Trans. Inf. & Syst. (Japanese Edition), vol.J80-D-II, no.5, pp.1192–1201, May 1997.

[20] Y. Zhang, Z. Li, and Y. Chen, “Gradient-based fast decision for intra prediction in HEVC,” Proc. IEEE Visual Commun. and Image Processing, San Diego, USA, pp.1–6, Nov. 2012.

[21] J. Kim, J. Yang, H. Lee, and B. Jeon, “Fast intra mode decision of HEVC based on hierarchical structure,” Proc. 8th Int’l Conf. on Information, Commun. and Signal Processing, Singapore, pp.1–4, Dec. 2011.

[22] Y. Gan, X. Zhao, and Q. Zhang, “Fast algorithm with early termination CU split and mode decision,” Int’l J. of Control and Automation, vol.8, no.2, pp.83–94, Feb. 2015.

[23] L. Shen, Z. Zhang, and P. An, “Fast CU size decision and mode decision algorithm for HEVC intra coding,” IEEE Trans. Consumer Electron., vol.59, no.1, pp.207–213, Feb. 2013.

Chihiro Tsutake received the B.E. from University of Fukui, Japan, in 2015. He is currently a master course student of Graduate school of Engineering, University of Fukui, where he is doing a research on video coding.

Yutaka Nakano received the B.E., M.E., and D.E. all in electronics from University of Fukui, in 1975, 1977, and 1999, respectively. He is currently an associate Professor in Dept. of Information Science, Graduate School of Eng., University of Fukui, Japan. His research interests include image and video processing.

Toshiyuki Yoshida received the B.E., M.E., and D.E. all in electronics from Tokyo Institute of Technology, Japan, in 1986, 1988, 1991. He was an associate professor in Tokyo Institute of Technology in 1994. He has moved to University of Fukui in 2003, where he is currently a professor in Dept. of Information Science. His research interests covers the image/video and signal processing and 3-D imaging. He is a co-author of several books, e.g., “Image Coding” (in Japanese, Corona Publishing) and “Image Information Coding” (in Japanese Ohmsha).