# Relative SATD-based Minimum Risk Bayesian Framework for Fast Intra Decision of HEVC

**Daehyeok Gwon and Haechul Choi**
The Department of Multimedia Engineering,
Hanbat National University, Daejeon, Korea
[e-mail: skyeye0530@naver.com, and choihc@hanbat.ac.kr]
*Corresponding author : Haechul Choi

## Abstract

High Efficiency Video Coding (HEVC) enables significantly improved compression performance relative to existing standards. However, the advance also requires high computational complexity. To accelerate the intra prediction mode decision, a minimum risk Bayesian classification framework is introduced. The classifier selects a small number of candidate modes to be evaluated by a rate-distortion optimization process using the sum of absolute Hadamard transformed difference (SATD). Moreover, the proposed method provides a loss factor that is a good trade-off model between computational complexity and coding efficiency. Experimental results show that the proposed method achieves a 31.54% average reduction in the encoding run time with a negligible coding loss of 0.93% BD-rate relative to HEVC test model 16.6 for the Intra_Main common test condition.

*Keywords:* HEVC, intra coding, fast encoder, intra mode decision.

## 1. Introduction

With the increasing demand for high-quality and high-resolution video content, High Efficiency Video Coding (HEVC) was developed by the ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations, working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) [1]–[4]. HEVC, the most recent international video coding standard, enables significantly improved coding efficiency compared to H.264/AVC [5]–[7] in the range of 50% bit rate reduction for equal perceptual video quality [8]. However, HEVC encoders are also expected to have higher complexity than H.264/AVC encoders [9]. Therefore, fast encoding algorithms have been broadly researched to decrease the encoding complexity of HEVC for real-time video encoding systems or power-constrained mobile devices.

Many advanced coding tools have been newly adopted in HEVC. In particular, a hierarchical coding structure plays an important role in the improved coding efficiency. It provides a highly flexible hierarchy of unit representation that includes three block concepts: coding unit (CU), prediction unit (PU), and transform unit (TU). The coding tree unit (CTU) is the core of the coding layer, which is analogous to the macroblock of previous standards. It consists of a luma coding tree block (CTB) and the corresponding chroma CTBs. The size $L \times L$ of a luma CTB can be chosen as $L = 16, 32$, or 64 samples. HEVC then supports a partitioning of the CTBs into smaller CUs using a tree structure and quadtree-like signaling [10]. The decision to use either inter or intra prediction to code a picture area is made at the CU level. The PU is the elementary unit for prediction and is defined after the last level of CU splitting. The PU size at which the intra prediction mode is established is the same as the CU size for all block sizes except for the smallest CU size. The TU is another transform and quantization-related unit whose size does not exceed that of the CU. In the HEVC intra prediction, the number of prediction modes for square PU sizes from $4 \times 4$ up to $32 \times 32$ increases to 35 modes: directional prediction with 33 different directional orientations, planar prediction (assuming an amplitude surface with a horizontal and vertical slope derived from PU boundaries), and DC prediction (a flat surface with a value matching the mean value of the boundary) [11]. These block concepts are helpful for optimizing each unit. However, HEVC encoders need to exhaust all the combinations of CU, PU, and TU to find the optimal solution, which is a very time-consuming process. Moreover, 35 intra prediction modes need to be evaluated for each PU to select the best direction. The increased number of intra prediction modes also requires substantial computational complexity.

To relieve the burden of intra prediction coding on encoders, the HEVC Test Model (HM) [12] first determines a few best candidate modes among all 35 intra prediction modes according to the sum of the absolute Hadamard transformed difference (SATD) and the number of bits required for the prediction mode. This process is called rough mode

decision (RMD) [13]. Then, by considering the strong correlation among spatially neighboring blocks, most probable modes (MPMs) are added to the candidate mode set. Finally, a full rate-distortion optimization (RDO) process is carried out only for the candidate mode set. Furthermore, a large amount of research effort has focused on speeding up the intra prediction mode decision. One category of this research is based on correlation among prediction directions of the candidate mode set, the spatially neighboring blocks, or previous-depth CUs [15]–[19]. The other category utilizes the gradient or edge of the current encoding block [20]–[23]. Most of the conventional research directly employs the candidate modes selected by the RMD process, but rarely considers the characteristics of RMD costs of the candidate modes. In experimental results, it is observed that a mode with far lower RMD cost than other modes is likely to be the best mode. Therefore, discrimination using the relative RMD costs would be useful to reduce the number of candidate modes further.

This paper introduces a modified minimum risk Bayesian classification framework employing a relative SATD that indicates how different SATDs are in a set of intra prediction modes. The classifier selects a small number of candidate modes, thereby reducing complexity. In experiments, to validate the performance of the proposed method, the proposed Bayesian classifier is compared with Lee's method [15] and Zhang's method [16]. Moreover, to further reduce the computational complexity, a combined scheme using Zhang's method, Lee's method, and the proposed method is also introduced.

The rest of this paper is organized as follows. In Section II, the conventional fast intra prediction mode decision methods are explained. In Section III, the relative SATD cost is defined. Then, in Section IV, a fast intra mode decision algorithm based on the minimum risk Bayesian classification is proposed. Section V compares the performance of the proposed algorithm with conventional work, and Section VI concludes this paper.

## 2. Related Work

The HM [12] adopts a fast intra mode decision method to reduce the complexity burden on HEVC encoders. This fast method consists of two phases. In the first phase, the N most promising candidate modes are selected from among all 35 intra prediction modes by the RMD process, where N is set to {8, 8, 3, 3, 3} for PUs of size $4 \times 4$, $8 \times 8$, $16 \times 16$, $32 \times 32$, and $64 \times 64$, respectively. In the RMD process, all intra prediction modes are evaluated with respect to the following cost function:

$$C = D_{Had} + \lambda \cdot R_{mode} \tag{1}$$

where D_Had is the SATD, R_mode represents the number of bits for a prediction mode, and C is the RMD cost. Then, by cons+0idering the strong correlation among the neighboring blocks, MPMs are added to the candidate mode set [14]. In the second phase, the candidate modes are thoroughly evaluated in the sense of the RDO. The prediction mode with the minimum rate-distortion (RD) cost is selected as the final prediction mode

[11]. Since the RD cost is calculated through the reconstruction of residual signal and bit estimation, including intra prediction, transformation, quantization, and entropy coding, the RDO process is computationally expensive. As mentioned above, the RMD process reduces the number of candidate modes evaluated in the RDO process, thereby reducing complexity. However, the RMD process has a fixed strategy for blocks with the same size, and it does not consider their different characteristics. If the number of candidate modes is more adaptable to the block characteristics, the complexity may be further reduced.

Recently, a number of fast algorithms have been proposed to reduce the computational complexity of the HEVC intra mode decision. Many of the approaches have been based on the similarity among intra prediction directions of the candidate mode set, the spatially adjacent blocks, or previous-depth CUs [15]–[19]. Lee et al. [15] proposed a simple, fast intra prediction mode decision method. In their work, if a mode with the minimal SATD cost is one of the MPMs, the mode is directly decided as being the best mode instead of performing the RDO process. Zhang et al. [16] introduced an adaptive, fast intra mode decision method, which finds a mode with the minimum SATD cost and employs the directional similarity between the minimum mode and the others in the candidate set to adaptively reduce the candidate modes for the RDO process. Zhao et al. [17] also reduced the candidate modes for the RDO process, utilizing direction information of the spatially adjacent CUs. Quanhe et al. [18] proposed an intra prediction mode decision strategy that sets a proper order of prediction modes according to the nearness of their directions and the characteristics of adjacent blocks. A fast intra mode decision based on a hierarchical structure [19] terminates the mode decision procedure using the intra prediction mode of the corresponding PU at the previous depth level and the size of TU at the current depth level. Shen et al

Other approaches have used gradient or edge characteristics since the intra coding is inherently based on predicting the directional structures that are present in typical video and image content [20]–[28]. Jiang et al. [20] proposed a gradient-based fast mode decision algorithm that calculates gradient directions and generates a gradient-mode histogram for each CU. Based on the distribution of the histogram, only a few of the intra prediction modes are chosen for the RMD and RDO processes. Another gradient-based algorithm [21] obtained the texture complexity and direction of CUs by applying intensity gradient filters and then excluded some of the intra prediction modes according to the texture direction in the mode decision process. An edge-based fast intra mode decision method [22] determined the candidate modes by analyzing the textures of the source image block. Considering the difference between the prediction directions of neighboring blocks, a fixed-point arithmetic based edge detector was designed to improve the direction detection accuracy. Yan et al. [23] proposed a group-based fast intra mode decision method. First, an early termination method is applied if the RMD cost of one mode is much smaller than that of the others. Second, the candidate modes are grouped together according to their angles, and then a pixel-based edge detection algorithm is applied to select the optimal angular direction. An edge-based CU size decision method decided split or non-split for CU according to complexities of global and local edge [24]. H. Zhang et al. [25] reduced candidates of rate-distortion optimized quantization based on SATD with 2:1

down-sampled prediction residual. Y. Shi et al. [26] determined PU level quadtree depth based on correlation of PU quadtree structure between current largest coding unit and its spatial and temporal neighbors. G. Tian et al. [27] utilized down-sampled texture complexity and prediction unit size of neighboring blocks to remove unnecessary operation for PU split. Shen et al. [28] restricted CU depth according to characteristic of image and skipped some prediction modes which were rarely used in CUs of the upper depth levels or spatially neighboring CUs.

On the other hand, various approaches have applied early decision schemes of coding depth based on temporally or spatially neighboring information [29]-[34]. Tao Fan et al. [29] proposed a fast CU size decision algorithm considering depth levels of neighboring CUs, distribution of rate distortion values and distribution of residual data. Lei Feng et al. [30] introduced a fast PU selection method that builds a saliency map for each largest coding block. Zhenglong et al. [31] determined CU depths according to a progressive gradient accumulation strategy obtained by adding up all the differences of every adjacent column and row. Liquan Shen et al. [32] proposed an early CU size decision algorithm based on the texture homogeneity. Xingang Liu et al. [33] also introduced a fast CU size decision algorithm that utilizes a CU complexity classifier built by using machine learning technology. H. Zhang et al. [34] proposed a fast mode decision that was made by priority classification according to coding information of spatial and temporal neighboring PUs. Liquan Shen et al. [35] proposed an intra mode classification and an early termination method of CU split according to intra mode information.

Most of the aforementioned fast intra prediction methods have been developed on the top of the HM that includes the RMD process. That is, those methods have taken account of a subset of intra prediction modes selected by the RMD process and tried to reduce the number of candidate modes for the RDO process. Note that those methods have focused only on the modes that are selected by the RMD process, but they have hardly considered the RMD costs. It is a fact that the RMD cost is not firmly reliable in the sense of the RDO because it utilizes SATD instead of the kinds of discrete cosine transform (DCT) or discrete sine transform (DST) used in HEVC. However, it is observed in experiments that the RMD costs roughly follow the same trend as the RDO costs, and a mode with substantially low RMD cost compared with other modes is likely to be a promising competitor for the optimal mode. Therefore, the RMD costs can be useful in the reduction of the candidate modes. In this paper, considering this observation, a fast intra prediction mode decision method is introduced, which takes full advantage of the RMD costs to further reduce the encoding complexity.
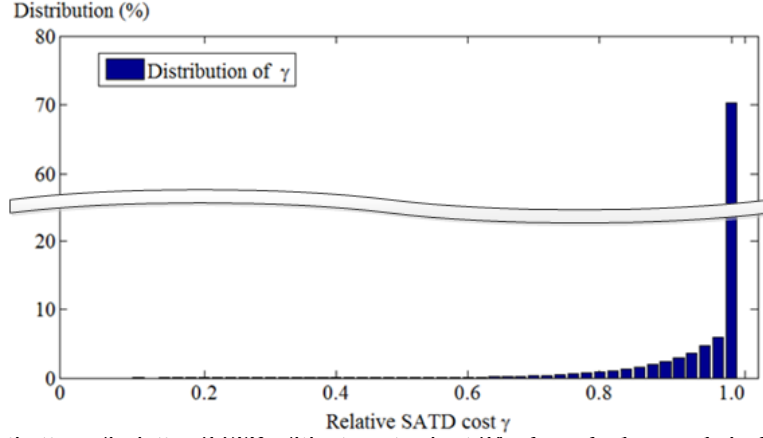
## 3. Relative SATD Cost

As described in Section II, the RMD process of the HM decides the first minimum cost N modes as the candidate set. In this paper, for clear explanation, RDO candidate set represents a subset of modes to be evaluated in the RDO process, and RMD modes indicates the modes selected by the RMD process. In the HM, the RMD modes are directly

used as the RDO candidate set. The total number, N, of the RMD modes depends only on the size of the PUs. For example, N is three for PUs greater than or equal to $16 \times 16$ and is eight for all other PUs. The RMD process constantly selects the fixed number of modes according the ranking of the RMD costs, regardless of how high the RMD costs are. If some of the RMD modes have substantially higher RMD costs than others, they are unlikely to be the best mode in the sense of RDO. Therefore, the modes with relatively higher RMD costs may induce unnecessary computational complexity in the RDO process. To solve this problem, the relatively higher RMD cost modes can simply be discarded by means of a predefined and absolute threshold. However, the SATD of the RMD cost as defined in (1) tends to vary widely according to the content of the current encoding block. For example, if the content is simple, the intra prediction is usually precise and SATD is therefore low. On the other hand, if the content is complex, the prediction may generate many residuals, and SATD thereby reaches a very high value. Therefore, it is very difficult to find an appropriate constant threshold value.

To efficiently remove the relatively higher RMD cost modes from the candidate set, the proposed method defines subsets of the RMD modes and selects an appropriate subset according to the SATD characteristics of the modes. Let $S_k$, $(1 \leq k \leq N)$, where N is the number of modes selected by the RMD process, be subsets of the RMD modes. $S_k$ has $k$ modes. If the sizes of the PUs are $4 \times 4$ or $8 \times 8$, $S_8$ consists of 8 modes selected by the RMD process, $m_1$ through $m_8$, where $m_i$ notes an intra prediction mode with the i-th minimum SATD, $(1 \leq i \leq k, \ m_i \in S_k)$. For all i less than j, the SATD of $m_i$ is less than or equal to that of $m_j$, and $m_1$ and $m_k$ are the minimum and maximum SATD modes in the subset $S_k$, respectively. $S_{k-1}$ can be defined by discarding the maximum SATD cost mode in $S_k$. In this way, other subsets, $S_1$ through $S_{k-2}$, are iteratively defined. For PUs greater than or equal to $16 \times 16$, the RMD process generates $S_3$ with $m_1, m_2$, and $m_3$.

By using the aforementioned subsets, the problem of reducing the number of RDO candidate modes can be converted into that of selecting a subset among $S_1$ through $S_N$ as the RDO candidate set. Deciding a subset $S_k$, of which $k$ is smaller than $N$, reduces the computational complexity since modes that do not belong to $S_k$, $m_{k+1}$ through $m_N$, are not input to the RDO process. For example, if $S_2$ is selected, only two modes, $m_1$ and $m_2$, are evaluated in the RDO process instead of all N modes. Selecting a subset that has fewer modes can thus reduce the complexity. However, an appropriate subset $S_k$ that maintains coding efficiency should be selected as the RDO candidate set. That is, the appropriate subset has to include the best mode in the sense of RDO, which is called the RDO best mode, $\hat{m}$. For example, when $\hat{m}$ has the $b$-th minimum SATD among $N$ RMD modes, the best subset that reduces complexity maximally while keeping coding efficiency is $S_k$ with $k$ equal to $b$. If the selected $k$ is less than $b$, computational complexity can be further reduced, but coding efficiency can also be reduced because $\hat{m}$ is discarded from the RDO candidate set. On the other hand, if the selected $k$ is larger than $b$, unnecessary computations for modes $m_j$, $(b < j \leq k)$, should be carried out in the RDO process. Like $m_j$, the mode that has a higher value of SATD than $\hat{m}$ is called the unnecessary higher-SATD mode, $\tilde{m}$, in this paper. $\tilde{m}$ causes unnecessary computation in the RDO

**Fig. 1.** Distribution of relative SATD of the target subset $\hat{S}$, where the last mode is the RDO best mode $\hat{m}$.

process. In conclusion, the proposed method is aimed at finding the subset $S_b$ that includes $\hat{m}$ and does not include any $\tilde{m}$. This subset $S_b$ is called the target subset, $\hat{S}$, in this paper.

To achieve the aim, the relative SATD of a subset $S_k$, $\gamma$, is defined by

$$\gamma = SATD_{m_1}/SATD_{m_k} \qquad (2)$$

where $SATD_{m_1}$ and $SATD_{m_k}$ are the SATDs of $m_1$ and $m_k$ in the subset $S_k$, respectively. For all $k$, $(1 \leq k \leq N)$, $\gamma$ of $S_k$ is normalized in the range between 0 and 1, and $\gamma$ of $S_{k-1}$ is larger than or equal to one of $S_k$. The $\gamma$ implies a ratio of the minimum SATD to the maximum SATD in a subset. By means of $\gamma$, it can be assessed how much higher the SATD of $m_k$ is than the minimum SATD. If $\gamma$ is closer to 1, all modes of $S_k$ have very similar SATD to the minimum SATD, $SATD_{m_1}$. On the contrary, if it is closer to 0, at least one mode including $m_k$ has a much higher SATD than the minimum SATD.

To exploit a relation between the target subset $\hat{S}$ and the relative SATD $\gamma$, experiments have been performed under common HM test conditions [36] for JCT-VC test sequences, including SteamLocomotiveTrain, Traffic, Cactus, ParkScene, BQMall, PartyScene, BlowingBubbles, and BQSquare. **Fig. 1** illustrates the distribution of the relative SATD $\gamma$ of the target subset $\hat{S}$, where the last mode is the RDO best mode $\hat{m}$. Note that the distribution of $\gamma$ of $\hat{S}$ equal to 1 is about 70%. In addition, for $\gamma$ of $\hat{S}$ that is higher than 0.9, the cumulative distribution is over 90%. As shown in **Fig. 1**, the target subset has a high correlation with the relative SATD. Thus, as $\gamma$ of a subset is closer to 1, the subset is highly likely to be the target subset.

## 4. Proposed Fast Intra Prediction Mode Decision

### 4.1 Minimum-risk Bayesian Classifier

To select a subset $S_k$ according to the relative SATD, $\gamma$, a modified Bayesian classification framework is introduced. First, the proposed method defines two-class $W = \{\omega_p, \omega_q\}$, where $\omega_p$ and $\omega_q$ stand for whether a subset $S_k$ includes the RDO best mode $\hat{m}$ or not, respectively. For instance, assuming that there are eight RMD modes and the RDO best mode $\hat{m}$ has the 5-th minimum SATD, subsets $S_5$, $S_6$, $S_7$, and $S_8$ correspond to $\omega_p$, and the others correspond to $\omega_q$. As mentioned in Section III, the proposed method is aimed at finding the target subset $\hat{S}$ that includes $\hat{m}$ and has the minimum number of modes. Thus, in case of the above instance, $\hat{S}$ is $S_5$.

Let $P(\omega_i|\gamma), i \in \{p, q\}$ be the a posteriori probability that $S_k$ belongs to $\omega_i$ given $\gamma$ of $S_k$. Here, action $\alpha_i$ corresponds to deciding that the true state of nature is $\omega_i$. For simple notation, let $\bar{C}_{i,j} = \bar{C}(\alpha_i|\omega_j), (i, j \in \{p, q\})$ be the loss incurred for deciding $\omega_i$ when the true state of nature is $\omega_j$. The loss can mean coding loss and complexity increase caused by a decision of the classifier. The loss can be assumed to be equal to 0 when the decision is true, i.e., $\bar{C}_{p,p} = \bar{C}_{q,q} = 0$. Therefore, the conditional risk associated with the decision $\omega_i$, $R(\alpha_i|\gamma)$, which is an expected loss, is represented as follows

$$R(\alpha_p|\gamma) = \bar{C}_{p,p}P(\omega_p|\gamma) + \bar{C}_{p,q}P(\omega_q|\gamma) = \bar{C}_{p,q}P(\omega_q|\gamma) \qquad (3)$$
$$R(\alpha_q|\gamma) = \bar{C}_{q,p}P(\omega_p|\gamma) + \bar{C}_{q,q}P(\omega_q|\gamma) = \bar{C}_{q,p}P(\omega_p|\gamma). \qquad (4)$$

There are various ways of expressing the minimum-risk decision rule, each having its own minor advantages. The fundamental rule is to decide $\omega_p$ if $R(\alpha_p|\gamma) < R(\alpha_q|\gamma)$. In term of the posterior probabilities, $\omega_p$ is decided if

$$\bar{C}_{p,q}P(\omega_q|\gamma) < \bar{C}_{q,p}P(\omega_p|\gamma). \qquad (5)$$

Meanwhile, the posteriori $P(\omega_i|\gamma)$ is given by Bayes rule [37] as

$$P(\omega_i|\gamma) = p(\gamma|\omega_i) \cdot P(\omega_i)/p(\gamma), \qquad (6)$$
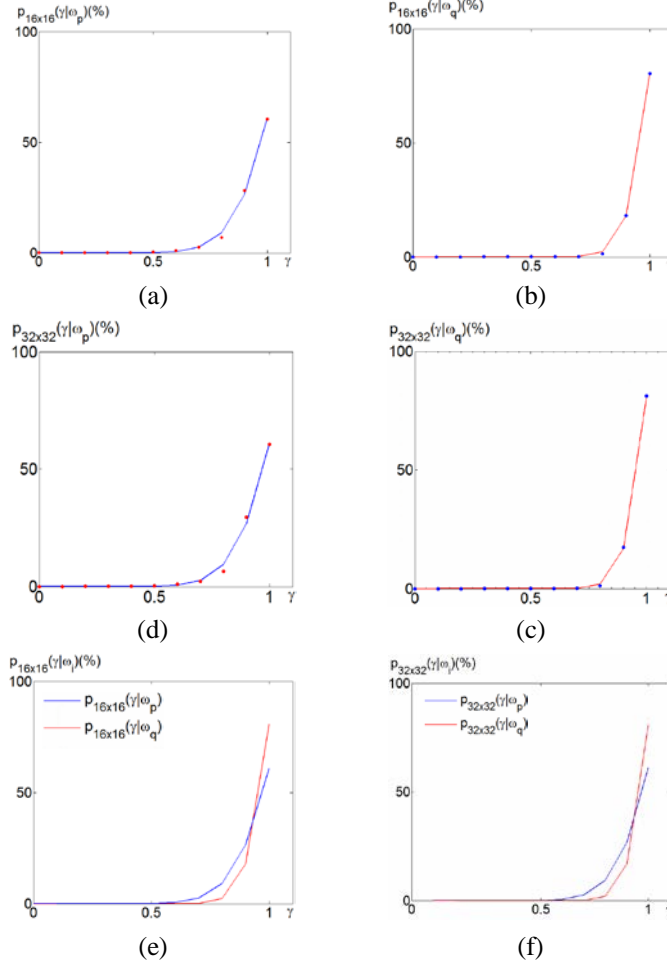
where $p(\gamma|\omega_i)$ is the conditional probability density function of $\gamma$ whose distribution depends on $\omega_i$, and $P(\omega_i)$ is the a priori probability. By using (6), the decision rule is instead expressed as follows

$$\bar{C}_{p,q}p(\gamma|\omega_q) \cdot P(\omega_q) < \bar{C}_{q,p}p(\gamma|\omega_p) \cdot P(\omega_p). \qquad (7)$$

Another alternative is to decide $\omega_1$ if

$$\frac{p(\gamma|\omega_p)}{p(\gamma|\omega_q)} > \frac{\bar{C}_{p,q}}{\bar{C}_{q,p}} \cdot \frac{P(\omega_q)}{P(\omega_p)}. \qquad (8)$$



**Fig. 2.** Actual distribution and modeled Gaussian distribution of $p(\gamma|\omega_i)$, (a) $p(\gamma|\omega_p)$ for $16 \times 16$ PUs, (b) $p(\gamma|\omega_q)$ for $16 \times 16$ PUs, (c) $p(\gamma|\omega_p)$ for $32 \times 32$ PUs, (d) $p(\gamma|\omega_q)$ for $32 \times 32$ PUs, (e) modeled Gaussian distributions of $p(\gamma|\omega_i)$ for $16 \times 16$ PUs, (f) modeled Gaussian distributions of $p(\gamma|\omega_i)$ for $32 \times 32$ PUs.

For simple notation, by substituting the loss term $\bar{C}_{p,q}/\bar{C}_{q,p}$ with loss factor $\lambda$ and a priori terms $P(\omega_q)/P(\omega_p)$ with $K$, the minimum-risk Bayesian classifier $\Psi(\gamma)$ can be finally written as follows

$$\Psi(\gamma) = \begin{cases} \omega_p & if \ \dfrac{p(\gamma|\omega_p)}{p(\gamma|\omega_q)} > \lambda \cdot K \\ \omega_q & otherwise \end{cases}. \tag{9}$$

According to (9), subsets including the RDO best mode $\hat{m}$ can be classified as $\omega_p$ and then a subset having the smallest number of modes is selected as the RDO candidate set from the subsets classified as $\omega_p$.

In (9), the conditional probability density function $p(\gamma|\omega_i)$ and a priori term $K$ can be derived through a training process. To obtain the statistical parameters, experiments have been carried out for JCT-VC test sequences [36]: SteamLocomotiveTrain, Traffic, Cactus, ParkScene, BQMall, PartyScene, BlowingBubbles, and BQSquare. To observe the probability distribution more strictly, $p(\gamma|\omega_i)$ is measured for every PU size (L × L), which is rewritten as $p_L(\gamma|\omega_i)$. Based on experimental results, $p_L(\gamma|\omega_i)$ is assumed to follow Gaussian distribution determined by the mean $\mu_{L,i}$ and the variance $\sigma_{L,i}{}^2$ such as

$$p_L(\gamma|\omega_i) = \frac{1}{\sqrt{2\pi}\sigma_{L,i}} \exp\left[-\frac{1}{2}\left(\frac{\gamma-\mu_{L,i}}{\sigma_{L,i}}\right)^2\right]. \tag{10}$$

**Fig. 2** shows the actual distribution and modeled Gaussian distribution of $p(\gamma|\omega_i)$ for 16 × 16 and 32 × 32 PUs given that the pattern is in class $\omega_i$. As observed in the figure, $p_{16}(\gamma|\omega_p)$ is greater than $p_{16}(\gamma|\omega_q)$ when the relative SATD $\gamma$ is smaller than about 0.9. Assuming that $\lambda \cdot K$ of (9) is equal to 1, the Bayesian decision boundary is set to $\gamma = 0.9$. In this decision boundary case, if $\gamma$ of a subset $S_k$ is smaller than 0.9, the subset is decided as $\omega_p$ by the minimum-risk Bayesian classifier, which means that the subset contains the RDO best mode $\hat{m}$. As $k$ of $S_k$ is higher, the RMD mode $m_k$ has a higher SATD and the $\gamma$ of $S_k$ gets lower. That is, $k$ and $\gamma$ of $S_k$ are inversely proportional to each other. A higher $k$ means that $S_k$ includes more modes. Therefore, a subset is likely to have the RDO best mode if its relative SATD $\gamma$ is low. However, the subset is likely to contain more of the unnecessary high-SATD modes as $k$ gets higher. In conclusion, it is desired to find the subset that has the minimum number of modes as well as the RDO best mode $\hat{m}$.

## 4.2 Fast Intra Prediction Mode Decision

A fast intra prediction algorithm using the minimum-risk Bayesian classifier with the relative SATD is proposed. As shown in **Fig. 3**, the proposed method consists of three steps. First, like the HM, the RMD process decides the first minimum cost N modes as the RMD modes where the N depends on PU sizes. Second, for subset $S_k$, starting from $S_1$ through $S_N$, its relative SATD $\gamma$ is calculated. Then the proposed minimum-risk Bayesian classifier using its relative SATD, $\Psi(\gamma)$, is applied to find the subset with the RDO best mode. If a subset $S_k$ is decided as $\omega_p$ by $\Psi(\gamma)$, this subset becomes the RDO candidate set $\acute{S}$ after unduplicated MPMs are added into the subset. In this case, modes from $m_{k+1}$ through $m_N$ are discarded from $\acute{S}$. Note that subsets are evaluated not in descending order from the

Rough Mode Decision (RMD)

8 for 4x4
8 for 8x8
3 for 16x16
3 for 32x32
3 for 64x64

Initialization : k = 1

$\gamma$ of $S_k$ Calculation

$k = k + 1$

$\Psi(\gamma) > \hat{\lambda} * K?$   no   →   $k \geq N - 1?$   no

yes                                        yes

$\acute{S} = S_k + MPM(s)$              $\acute{S} = S_N + MPM(s)$

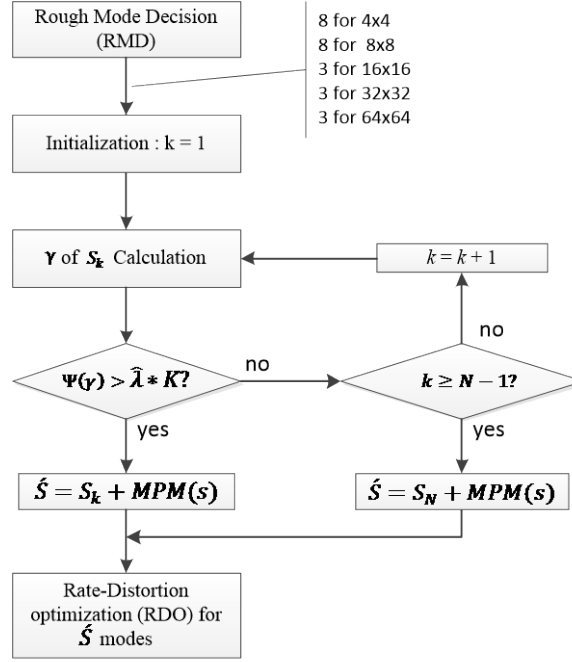Rate-Distortion optimization (RDO) for $\acute{S}$ modes

**Fig. 3.** The proposed fast intra prediciton mode decision algorithm.

subset $S_N$, but in ascending order from the subset $S_1$. This ascending order is for finding the subset with the minimum number of modes among subsets decided as $\omega_p$ by $\Psi(\gamma)$. Due to the ascending order evaluation, as many as possible of the unnecessary high-SATD modes are discarded from the RDO candidate set $\acute{S}$. Finally, the best intra prediction mode is determined through the conventional RDO process for the RDO candidate set $\acute{S}$ selected by the proposed method. The conventional HM performs the RDO process for all the RMD modes, whereas the proposed method does so for the subset $\acute{S}$, discarding some modes. This reduced number of RDO candidate modes can decrease the computational complexity.

## 4.3 Loss Factor Decision

The performance of the proposed method relies on the loss factor $\lambda$ equal to $\bar{C}_{p,q}/\bar{C}_{q,p}$ in (8). The loss factor is highly related to how costly the classifier mistakes are, and it can treat situations in which some kinds of classification mistakes are costlier than others. The proposed method focuses on the $\bar{C}_{p,q}$ that is the loss incurred for deciding $\omega_p$ when the true state of nature is $\omega_q$. When a subset does not include the RDO best mode, the wrong decision that the subset has the RDO best mode obviously results in coding loss. The higher the loss factor is, the more carefully the proposed classifier makes a decision. For example, when $\lambda$ is equal to 0, that means the mode discarding has no loss, so the subset $S_1$ including only $m_1$ is always selected by $\Psi(\gamma)$. In this case, computational complexity decreases
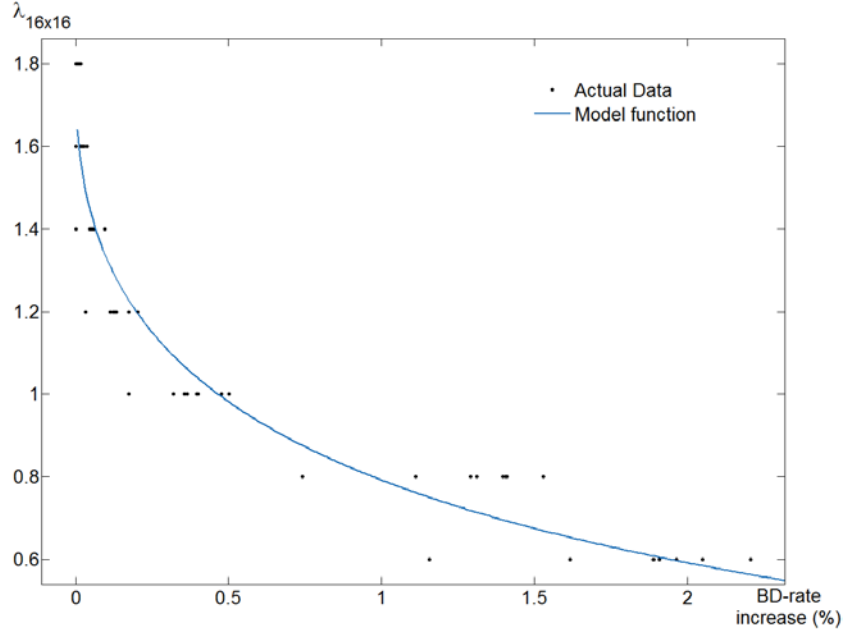
**Fig. 4.** BD-rate increase according to the loss factor and its exponential fitting model for PUs with $16 \times 16$

maximally, but coding loss increases. If $\lambda$ is an extremely high value, meaning that discarding the mode has a high risk, $\Psi(\gamma)$ always makes the decision as the $\omega_q$. Thus, the subset $S_N$ including all of the RMD modes is selected as the RDO candidate set just like HM. Therefore, coding efficiency is retained as well as the HM without any complexity reduction. As shown in the example cases, the loss factor provides a useful feature for controlling the tradeoff between coding efficiency and complexity reduction. To achieve complexity reduction while retaining an acceptable coding loss, the proposed Bayesian classifier has to set an appropriate value of $\lambda$.

To provide a guide for determining the loss factor $\lambda$, various loss factor values were simulated under the common HM test conditions with All Intra configuration for JCT-VC test sequences, including SteamLocomotiveTrain, Traffic, Cactus, ParkScene, BQMall, PartyScene, BlowingBubbles, and BQSquare [36]. Quantization parameters, 22, 27, 32, and 37, are used to calculate BD-rate. **Fig. 4** shows the coding loss when the proposed method is applied only to PUs with $16 \times 16$ relative to the HM according to various loss factor values. It is observed that the BD-rate [38] of coding loss monotonically decreases as the loss factor increases. The experimental results can be modeled by an exponential function. The modeled loss factor $\hat{\lambda}$ is represented as follows

$$\hat{\lambda} = \mu_1 \times e^{-\mu_2 * \beta^{\mu_3}}, \tag{11}$$

**Table 1.** Training and test sequences.

| Sequences | Training Sequences | Test Sequences |
|---|---|---|
| Class A | SteamLocomotiveTrain<br>Traffic | PeopleOnStreet<br>NebutaFestival |
| Class B | Cactus<br>ParkScene | Kimono<br>BasketballDrive<br>BQTerrace |
| Class C | BQMall<br>PartyScene | BasketballDrill<br>RaceHorses |
| Class D | BlowingBubbles<br>BQSquare | BasketballPass<br>RaceHorses |
| Class E |  | FourPeople<br>Johnny<br>KistenAndSara |
| Class F |  | BasketballDrillText<br>ChinaSpeed<br>SlideEditing<br>SlideShow |

where $\beta$ indicates the BD-rate increase and $\mu_1$, $\mu_2$, and $\mu_3$ are parameters of the exponential function to fit the experimental results to the model as closely as possible. To achieve a more accurate fitting function, for each PU size ($L \times L$), $\hat{\lambda}_{L \times L}$ and the parameters $\mu_{i,L}$ ($i \in \{1, 2, 3\}$) can be calculated. In **Fig. 4**, $\mu_{1,16}$, $\mu_{2,16}$, and $\mu_{3,16}$ are 1.784, 0.8113, and 0.4423, respectively, which minimizes the sum of squared errors. Given an acceptable coding loss represented by $\beta$, an appropriate loss factor can be derived easily by using the exponential model. In this way, the proposed method provides a useful feature for controlling the tradeoff between computational complexity and coding efficiency.

## 5. Experimental Results

To evaluate the performance of the proposed Bayesian classifier, it is compared with previous work by Lee [15] and Zhang [16], as well as HM [12]. All of these methods are implemented on the top of HM 16.6. Encoder controls follow the common HM test conditions with All Intra Main configuration [33]. In detail, the CTB indicating the maximum CU size has a fixed size of $64 \times 64$ pixels. The maximum depth of CTB is set to four, which allows CUs of sizes $8 \times 8$ to $64 \times 64$. The BD-rates, which provide the relative coding gain between two methods based on the average difference between their RD-curves, are commonly measured for objective quality evaluation by adopting HM as the anchor and utilizing quantization parameters of 22, 27, 32, and 37. JCT-VC test sequences including classes A, B, C, D, E, and F are used. Classes A, B, C, D, and E are camera captured content and class F contains screen content sequences. **Table 1** shows the training sequences and test sequences that are used for the parameter derivation of the Bayesian classifier and the performance evaluation, respectively. To compare the performance of computational complexity, encoding times are measured with encoding time saving (ETS) as follows

$$ETS(\%) = (ET_{ref} - ET_{test})/ET_{ref} \times 100, \qquad (12)$$

where $ET_{ref}$ and $ET_{test}$ are encoding times of HM 16.6 and a test method, respectively.

**Table 2.** Performances of the proposed method accodring to $\boldsymbol{\beta}$.

| Class | Sequence name | $\beta = 0.05$ | | $\beta = 0.2$ | | $\beta = 0.4$ | | $\beta = 0.6$ | | $\beta = 0.8$ | | $\beta = 1.0$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) |
| Class A | PeopleOnStreet | 0.31 | 17.27 | 0.62 | 22.72 | 0.89 | 25.23 | 1.12 | 26.98 | 1.30 | 27.87 | 1.45 | 29.13 |
| | NebutaFestival | 0.09 | 14.75 | 0.20 | 21.56 | 0.27 | 25.35 | 0.32 | 27.94 | 0.36 | 29.34 | 0.39 | 30.71 |
| Class B | Kimono | 0.15 | 15.7 | 0.37 | 21.79 | 0.50 | 24.7 | 0.59 | 26.63 | 0.67 | 27.56 | 0.72 | 28.74 |
| | BasketballDrive | 0.13 | 16.69 | 0.38 | 22.42 | 0.58 | 25.53 | 0.79 | 27.55 | 0.96 | 28.39 | 1.10 | 29.64 |
| | BQTerrace | 0.08 | 16.29 | 0.23 | 22.38 | 0.37 | 25.76 | 0.52 | 28.01 | 0.66 | 29.14 | 0.77 | 30.6 |
| Class C | BasketballDrill | 0.11 | 16.46 | 0.32 | 22.11 | 0.56 | 24.83 | 0.76 | 26.55 | 0.96 | 27.47 | 1.14 | 28.6 |
| | RaceHorses | 0.14 | 13.99 | 0.36 | 20.59 | 0.55 | 23.99 | 0.74 | 26.19 | 0.90 | 27.36 | 1.04 | 28.87 |
| Class D | BasketballPass | 0.22 | 16.4 | 0.52 | 22.51 | 0.84 | 25.13 | 1.09 | 27.05 | 1.34 | 27.96 | 1.52 | 29.2 |
| | RaceHorses | 0.21 | 13.92 | 0.49 | 20.60 | 0.78 | 23.97 | 1.04 | 26.35 | 1.27 | 27.74 | 1.46 | 29.25 |
| Class E | FourPeople | 0.23 | 19.60 | 0.53 | 23.38 | 0.77 | 25.72 | 0.99 | 28.86 | 1.16 | 28.56 | 1.31 | 28.80 |
| | Johnny | 0.23 | 20.52 | 0.62 | 24.05 | 0.82 | 26.33 | 1.02 | 29.60 | 1.17 | 29.26 | 1.30 | 29.45 |
| | KristenAndSara | 0.26 | 20.49 | 0.60 | 23.14 | 0.84 | 25.39 | 1.05 | 29.96 | 1.22 | 28.18 | 1.34 | 28.59 |
| **Average (Class A ~ Class E)** | | **0.18** | **16.84** | **0.44** | **22.27** | **0.65** | **25.16** | **0.84** | **27.64** | **1.00** | **28.24** | **1.13** | **29.30** |
| Class F (Screen Contents) | BasketballDrillText | 0.25 | 17.56 | 0.53 | 21.67 | 0.78 | 24.07 | 1.01 | 27.83 | 1.21 | 27.84 | 1.40 | 28.33 |
| | ChinaSpeed | 0.68 | 17.92 | 1.44 | 22.77 | 2.10 | 24.88 | 2.61 | 27.91 | 3.04 | 27.86 | 3.40 | 28.56 |
| | SlideEditing | 1.11 | 22.75 | 2.01 | 25.30 | 2.77 | 26.91 | 3.32 | 30.04 | 3.78 | 29.91 | 4.18 | 30.48 |
| | SlideShow | 0.71 | 27.88 | 1.12 | 27.94 | 1.54 | 28.69 | 1.73 | 31.61 | 2.03 | 29.79 | 2.16 | 30.02 |
| **Overall Average** | | **0.31** | **18.01** | **0.65** | **22.81** | **0.94** | **25.41** | **1.17** | **28.07** | **1.38** | **28.39** | **1.54** | **29.31** |

**Table 3.** Performances of Lee's method [15] and Zhang's method [16].

| Class | Sequence name | Lee [15] | | Zhang [16] | |
|---|---|---|---|---|---|
| | | BD-BR (%) | ETS (%) | BD-BR (%) | ETS (%) |
| Class A | PeopleOnStreet | 0.86 | 21.43 | 0.17 | 15.68 |
| | NebutaFestival | 0.39 | 17.10 | 0.04 | 17.3 |
| Class B | Kimono | 1.26 | 23.26 | 0.05 | 15.37 |
| | BasketballDrive | 0.99 | 26.03 | 0.17 | 15.18 |
| | BQTerrace | 0.57 | 24.83 | 0.10 | 15.65 |
| Class C | BasketballDrill | 0.76 | 23.00 | 0.20 | 14.46 |
| | RaceHorses | 0.82 | 19.94 | 0.15 | 15.39 |
| Class D | BasketballPass | 0.94 | 22.28 | 0.30 | 15.09 |
| | RaceHorses | 0.70 | 21.04 | 0.27 | 15.67 |
| Class E | FourPeople | 1.05 | 25.31 | 0.34 | 16.15 |
| | Johny | 1.12 | 28.01 | 0.29 | 15.83 |
| | KristenAndSara | 1.04 | 27.65 | 0.35 | 15.51 |
| **Average (Class A ~ Class E)** | | **0.88** | **23.32** | **0.20** | **15.61** |
| Class F (Screen Contents) | BasketballDrillText | 0.82 | 23.68 | 0.40 | 17.38 |
| | ChinaSpeed | 2.34 | 24.30 | 1.23 | 16.33 |
| | SlideEditing | 3.31 | 28.32 | 1.59 | 17.82 |
| | SlideShow | 1.13 | 31.45 | 0.54 | 16.57 |
| **Overall Average** | | **1.13** | **24.23** | **0.39** | **15.96** |

First, the individual performance of the proposed fast intra mode decision method employing the modified minimum-risk Bayesian classifier with the relative SATD is evaluated. Table 2 shows the performance of the proposed method in terms of coding efficiency and computational complexity when the acceptable BD-rate increases $\beta$, which determines the loss factor $\hat{\lambda}$ of the proposed classifier like (11), is equal to 0.05, 0.2, 0.4, 0.6, 0.8, and 1.0. According to $\beta$, the proposed Bayesian classifier reduces complexity by 18.01% to 29.30% in the ETS with negligible coding loss of 0.31% to 1.13% in the BD-rate relative to HM 16.6 for JCT-VC test sequences including classes A, B, C, D, and E. The experimental results of class F, screen content sequences, have higher complexity reduction compare to other classes, but also have increased coding loss. It is observed that both coding loss and complexity reduction increase as $\beta$ is a higher value. The high $\beta$ tends to generate a low value of the loss factor $\hat{\lambda}$ as in (11), which implies that the risk of the mode discarding is low. Affected by the low loss factor, the proposed minimum-risk Bayesian classifier $\Psi(\gamma)$ is more likely to select subsets with small numbers of modes for the RDO candidate set. This is the reason for the higher coding loss and the higher complexity reduction when the $\beta$ is high. Even though the coding loss gets higher, it is just negligible at the maximum 1.13% BD-rate. The experimental results verify that the proposed method can reduce computational complexity significantly while retaining coding efficiency. The experiments also demonstrate that the computational complexity reduction and the coding efficiency are tightly associated with the loss factor determined by $\beta$. Thus, the proposed method can provide a good trade-off model between the computational complexity and the coding efficiency that is represented as an exponential function as in (11).

In most ranges of $\beta$, the sequence NebutaFestival has the largest encoding complexity reduction. For instance, its ETS is 30.71% when $\beta$ is equal to 1.0. Furthermore, this sequence has a very low coding loss. Because it has very sharp edges and dominant directions of texture, the SATD costs of intra prediction modes tend to be significantly different from each other, and the relative SATD cost of subsets thereby decreases rapidly as the total mode number of the subsets increases. Therefore, subsets with small numbers of modes are likely to be decided as the RDO candidate set by the proposed classifier. On the other hand, sequences PeopleOnStreet, BasketballPass, and RaceHorses containing many moving objects and motion blurs have worse coding losses since the RMD modes have similar SATD costs and thus it may be difficult to find the target subset $\hat{S}$ by using only the relative SATD.

For comparison with the conventional methods, experimental results of Lee's method [15] and Zhang's method [16] are listed in Table 3. Lee's method achieves an average encoding time reduction of 23.32% in the ETS with coding loss of an average 0.88% BD-rate relative to HM 16.6. Compared with Lee's method, the proposed Bayesian classifier can achieve a slightly better complexity reduction of 25.41% in the ETS with almost the same coding loss when $\beta$ is equal to 0.4. Zhang's method reduces the encoding time by an average 15.5% in the ETS while the coding loss is 0.20% BD-rate relative to HM 16.6. Compared with Zhang's method, the proposed Bayesian classifier also obtains
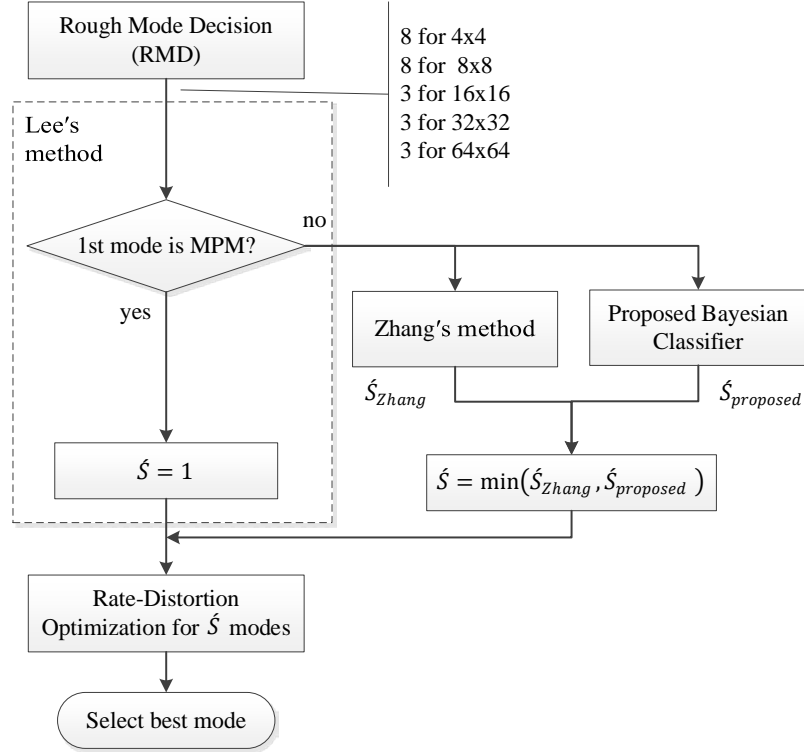
**Fig. 5.** The combined scheme with Zhang's method, Lee's method, and the proposed Bayesian classifier.

similar coding loss with increased the ETS when $\beta$ equal to 0.05. As shown in these experiments, Lee's method obtains more complexity reduction with lower coding efficiency, whereas Zhang's method achieves less complexity reduction with higher coding efficiency. It is clear that the complexity reduction and the coding efficiency are inversely proportional. Compared with these methods showing the constant performance, the proposed Bayesian classifier can provide not only the same or slightly better performance, but also a wide range of complexity reduction control from 16.84% to 29.30% in the ETS by controlling $\beta$.

The proposed method can be harmonized with Zhang's method and/or Lee's method. **Fig. 5** shows the combined scheme that harmonizes the proposed Bayesian classifier, Zhang's method, and Lee's method. In this combined scheme, Lee's method is first applied since it is very simple and the RDO candidate set of Lee's method is only one mode. If the criteria of Lee's method are not satisfied, the proposed Bayesian classifier and Zhang's method are conducted simultaneously. The proposed Bayesian classifier decides a preliminary RDO candidate set $\acute{S}_{proposed}$ by using decision function $\Psi(\gamma)$. Meanwhile, another preliminary RDO candidate set $\acute{S}_{Zhang}$ is obtained by Zhang's method. The subset

**Table 4.** Performances of combinations of the proposed method, Lee's method, and Zhang's method ($\beta = 0.05$).

| Class | Sequence name | Proposed + Lee[15] | | Proposed + Zhang[16] | | Proposed + Lee [15]+ Zhang[16] | |
|---|---|---|---|---|---|---|---|
| | | BD-rate Y (%) | ETS (%) | BD-rate Y (%) | ETS (%) | BD-rate Y (%) | ETS (%) |
| Class A | PeopleOnStreet | 0.98 | 28.89 | 0.41 | 21.50 | 1.06 | 31.54 |
| | NebutaFestival | 0.34 | 27.38 | 0.11 | 21.66 | 0.35 | 31.76 |
| Class B | Kimono | 0.77 | 29.03 | 0.18 | 20.08 | 0.79 | 31.57 |
| | BasketballDrive | 1.01 | 31.31 | 0.26 | 20.98 | 1.10 | 33.41 |
| | BQTerrace | 0.67 | 31.01 | 0.16 | 21.79 | 0.73 | 33.77 |
| Class C | BasketballDrill | 0.94 | 28.91 | 0.28 | 20.67 | 1.08 | 31.44 |
| | RaceHorses | 0.78 | 27.62 | 0.24 | 20.00 | 0.86 | 31.10 |
| Class D | BasketballPass | 1.18 | 28.98 | 0.43 | 20.63 | 1.31 | 31.87 |
| | RaceHorses | 1.04 | 27.19 | 0.41 | 20.01 | 1.16 | 30.78 |
| Class E | FourPeople | 0.74 | 26.57 | 0.46 | 22.77 | 0.92 | 30.25 |
| | Johny | 0.76 | 28.35 | 0.47 | 22.41 | 0.93 | 30.96 |
| | KristenAndSara | 0.73 | 26.70 | 0.49 | 21.79 | 0.94 | 29.86 |
| **Average (Class A ~ Class E)** | | **0.83** | **28.50** | **0.33** | **21.19** | **0.94** | **31.53** |
| Class F (Screen contents) | BasketballDrillText | 0.72 | 24.95 | 0.53 | 22.12 | 0.95 | 29.38 |
| | ChinaSpeed | 1.94 | 24.71 | 1.50 | 22.32 | 2.69 | 28.91 |
| | SlideEditing | 2.68 | 27.98 | 2.01 | 26.55 | 3.53 | 33.39 |
| | SlideShow | 1.32 | 30.60 | 0.99 | 27.39 | 1.48 | 32.49 |
| **Overall Average** | | **1.04** | **28.14** | **0.56** | **22.04** | **1.24** | **31.41** |

with fewer modes between $\acute{S}_{proposed}$ and $\acute{S}_{Zhang}$ is selected as the final RDO candidate set $\acute{S}$. If the number of modes of $\acute{S}_{proposed}$ and $\acute{S}_{Zhang}$ are identical, $\acute{S}_{proposed}$ is selected as the RDO candidate set. The individual combinations with Lee's method and Zhang's method, respectively, are also implemented. **Table 4** lists the performance of the various combinations relative to HM 16.6 under the condition of $\beta$ equal to 0.05. As shown in the table, each of the conventional methods achieves more than 4% complexity reduction in the ETS with very slight additional coding loss when it is combined with the proposed method. In particular, the combined scheme of all the three methods obtains an average 31.53% ETS with a coding loss of 0.94% BD-rate. This scheme has more complexity reduction with less coding loss than the individual performance of the proposed method at $\beta$ equal to 1.0.

All of Lee's method, Zhang's method, and the proposed method have an approach that speeds up intra prediction mode decision by reducing the number of intra prediction candidate modes in the RDO process. The proposed method may be harmonized with fast RDO methods having other approaches. For example, the proposed method can be combined with early CU splitting and pruning methods that decide the size of CU quickly. The CU pruning methods should also compute costs of all intra prediction modes for each CU size. Therefore, the proposed method reducing the number of the candidate modes is able to be harmonized with these methods. The harmonized method may be efficiently

faster than a single method like the results of **Table 4**.

## 6. Conclusion

This paper introduces a modified minimum-risk Bayesian classifier using the relative SATD to accelerate the intra prediction mode decision of HEVC encoders. The proposed Bayesian classifier decides the subset that is likely to have the RDO best mode as the RDO candidate set, which can reduce the total number of modes to be evaluated in the RDO process, discarding unnecessary higher-SATD modes. Furthermore, the proposed method provides a good trade-off model between the computational complexity and the coding efficiency that is represented by the loss factor. As shown in the experimental results, the proposed method could reduce the encoding time by up to 30% with a negligible coding loss of 1% BD-rate for the All Intra Main case. Moreover, the conventional methods, Lee's and Zhang's methods, could achieve more complexity reduction combined with the proposed method. The proposed method can contribute to the design of fast HEVC encoders. For future work, more encoding time saving would be achieved if the proposed method is combined with early CU splitting and pruning methods.

## Acknowledgments

## References

[1]  G. J. Sullivan and J.-R. Ohm, "Recent developments in standardization of High Efficiency Video Coding (HEVC)," in *Proc. of 33rd SPIE Appl. Dig. Image Process.*, vol. 7798, pp. 7798–7830, Aug. 2010. Article (CrossRef Link).

[2]  T. Wiegand et al., "Special section on the joint call for proposals on High Efficiency Video Coding (HEVC) standardization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 12, pp. 1661–1666, Dec. 2010. Article (CrossRef Link).

[3]  Joint Collaborative Team on Video Coding (JCT-VC), "High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Consent)," *JCTVC-L1003*, Geneva, Jan. 2013.

[4]  Gary J. Sullivan et al., "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 12, pp. 1649–1668, Sept. 2012. Article (CrossRef Link).

[5]  ITU-T and ISO/IEC JTC 1, "Advanced Video Coding for Generic Audiovisual Services," *ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC)*, version 1, 2003, version 2, 2004, versions 3, 4, 2005, versions 5, 6, 2006, versions 7, 8, 2007, versions 9, 10, 11, 2009, versions 12, 13, 2010, versions 14, 15, 2011, version 16, 2012.

[6]  G. J. Sullivan and T. Wiegand, "Video compression - from concepts to the H.264/AVC standard," in *Proc. of IEEE*, vol. 93. no. 1, pp. 18–31, Jan. 2005. Article (CrossRef Link).

[7]  T. Wiegand et al., "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003. Article (CrossRef Link).

[8]   J.R. Ohm, G. J. Sullivan, and H. Schwarz, "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, Oct. 2012. Article (CrossRef Link).

[9]   F. Bossen et al., "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, Dec. 2012. Article (CrossRef Link).

[10]  H. Samet, "The quadtree and related hierarchical data structures," *ACM Comput. Surveys*, vol. 16, issue 2, pp. 187–260, Jun. 1984. Article (CrossRef Link).

[11]  J. Lainema et al., "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2010. Article (CrossRef Link).

[12]  High Efficiency Video Coding Test Model Software 16.6, https://hevc.hhi.fraunhofer.de/svn/svnHEVCSoftware.

[13]  Y. Piao, J. Min, and J. Chen, "Encoder improvement of unified intra prediction," *JCTVC-C207*, Guangzhou, Oct. 2010.

[14]  L. Zhao et al., "Further Encoder Improvement of intra mode decision," *JCTVC-D283*, Daegu, Jan. 2011.

[15]  S. Lee, S. Park, and E. Jang, "Fast intra prediction mode decision based on rough mode decision and most probable mode in HEVC," *Journal of Broadcast Engineering (JBE)*, 19(2), pp.158–165, 2014. Article (CrossRef Link).

[16]  M. Zhang, C. Zhao, and J. Xu, "An Adaptive Fast Intra Mode Decision in HEVC," in *Proc. of 19th IEEE International Conference on Image Processing (ICIP)*, pp. 221–224, Oct. 2012. Article (CrossRef Link).

[17]  L. Zhao et al., "Fast mode decision algorithm for intra prediction in HEVC," in *Proc. of IEEE Vis. Commun. Image Process. (VCIP)*, pp. 1–4, Nov. 2011. Article (CrossRef Link).

[18]  Y. Quanhe, R. Yaocheng, and H. Yun, "Fast intra mode decision strategy for HEVC," in *Proc. of IEEE China Summit Int. Conf. Signal Inf. Process. (ChinaSIP)*, pp. 500–504, Jul. 2013. Article (CrossRef Link).

[19]  J. Kim et al., "Fast intra mode decision of HEVC based on hierarchical structure," in *Proc. of 8th Int. Conf. Inf., Commun. Signal Process. (ICICS)*, pp. 1–4, Dec. 2011. Article (CrossRef Link).

[20]  W. Jiang, H. Ma, and Y. Chen, "Gradient based fast mode decision algorithm for intra prediction in HEVC," in *Proc. of Int. Conf. Consum. Electron., Commun. Netw. (CECNet)*, pp. 1836–1840, Apr. 2012. Article (CrossRef Link).

[21]  Y. Zhang, Z. Li, and B. Li, "Gradient-based fast decision for intra prediction in HEVC," in *Proc. of IEEE Vis. Commun. Image Process. (VCIP)*, pp. 1–6, Nov. 2012. Article (CrossRef Link).

[22]  G. Chen et al., "Fast HEVC intra mode decision using matching edge detector and kernel density estimation alike histogram generation," in *Proc. of IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. 53–56, May. 2013. Article (CrossRef Link).

[23]  S. Yan et al., "Group-based fast mode decision algorithm for intra prediction in HEVC," in *Proc. of Int. Conf. Signal Image Technol. Internet Based Syst.*, pp. 225–229, Dec. 2012. Article (CrossRef Link).

[24]  B. Min et al. "A fast CU size decisioni algorithm for the HEVC intra encoder," *IEEE Transactions on Circuits and Systems for Video Technology*, issue 5, pp. 892-896, May, 2015, Article (CrossRef Link).

[25]  H. Zhang et al. "Fast intra prediction for high efficiency video coding," *Advances in Multimedia Information Processing. PCM 2012. Lecture Notes in Computer Science*, vol. 7674, 2012. Article (CrossRef Link).

[26] Y. Shi et al. "Content based fast prediction unit quadtree depth decision algorithm for HEVC," *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, May, 2013. Article (CrossRef Link).

[27] G. Tian et al. "Content adaptive prediction unit size decision algorithm for HEVC intra coding," *Picture Coding Symposium (PCS)*, May, 2012, Article (CrossRef Link).

[28] L. Shen et al. "Fast CU size decision and mode decision algorithm for HEVC intra coding," IEEE Transactions on Consumer Electronics, vol. 59, issue 1, pp.207-213, February, 2013, Article (CrossRef Link).

[29] Tao FAN, Guozhong WANG, Xiwu SHANG, "Fast Coding Unit Size Decision in HEVC Intra Coding," *IEICE Transactions on Information and Systems*, issue 7, pp. 1953-1956, 2016. Article (CrossRef Link).

[30] Lei Feng, Ming Dai, Chun-lei Zhao, Jing-ying Xiong, "Fast prediction unit selection method for HEVC intra prediction based on salient regions," *Optoelectronics Letters*, vol. 12, issue 4, pp. 316-320, July, 2016. Article (CrossRef Link).

[31] Zhenglong, Guozhong Wang, Tao Fan, Guowei Teng, "Fast Intra Encoding Decisions Based on Horizontal-Vertical Progressive Gradient Accumulation for HEVC," *International Forum of Digital TV and Wireless Multimedia Communication*, vol. 685, pp. 255-264, March, 2017. Article (CrossRef Link).

[32] Liquan Shen, Zhaoyang Zhang, Zhi Liu, "Effective CU Size Decision for HEVC Intra coding," *IEEE Transactions on Image Processing*, vol. 23, issue 10, pp. 4232-4241, October, 2014. Article (CrossRef Link).

[33] Xingang Liu, Yayong Li, Deyuan Liu, Peicheng Wang Laurence T. Yang, "An Adaptive CU Size Decision Algorithm for HEVC Intra Prediction based on Complexity Classification using Machine Learning," *IEEE Transactions on Circuits and Systems for Video Technology*, Issue 99, November, 2017. Article (CrossRef Link).

[34] H. Zhang et al. "Priority classification based fast intra mode decision for high efficiency video coding," *Picture Coding Symposium (PCS)*, 2013, Article (CrossRef Link).

[35] Liquan Shen, Zhaoyang Zhang, Zhi Liu, "Efficient Intra Mode Selection for Depth-Map Coding Utilizing Spatiotemporal, Inter-Component and Inter-View Correlations in 3D-HEVC," *(2018) IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4195-4206, September, 2018. Article (CrossRef Link).

[36] Joint Collaborative Team on Video Coding (JCT-VC), "Common HM test conditions and software reference configurations," *JCTVC-I1101*, Geneva, May. 2012.

[37] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification (2nd ed.), John Wiley & Sons, 2012.

[38] G. Bjontgaard, "Calculation of Average PSNR Differences between RD-Curves," *ITU-T SG16 Q.6 VCEG*, Doc. VCEG-M33, 2001.

**Daehyeok Gwon** received his B.S. and M.S. degree in multimedia engineering from Hanbat National University, Daejeon, Korea in 2012 and 2014, respectively. He is currently pursuing a Ph.D. degree in multimedia engineering at Hanbat National University. His current research interests include image processing, video coding, and video transmission.

**Haechul Choi** received his B.S. in electronics engineering from Kyungpook National University, Daegu, Korea, in 1997, and his M.S. and Ph.D. in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1999 and 2004, respectively. He is an associate professor in the Information and Communication Engineering at Hanbat National University, Daejeon, Korea. From 2004 to 2010, he was a Senior Member of the Research Staff in the Broadcasting Media Research Group of the Electronics and Telecommunications Research Institute (ETRI). His current research interests include image processing, video coding, and video transmission.