IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 11, 2019, accepted October 7, 2019, date of publication October 21, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947260

Efficient Hard-Decision Quantization Using an Adaptive Deadzone Offset Model for Video Coding

HAIBING YIN[1], (Member, IEEE), HONGKUI WANG[2], XIAOFENG HUANG[1], AND HUABING YIN[3]

[1]School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[2]School of Electronic Information and Communication, Huazhong University of Science and Technology, Wuhan 430000, China
[3]COSCO Shipping Specialized Carries Company Ltd., Guangzhou 510623, China

Corresponding author: Haibing Yin (yhb@hdu.edu.cn)

This work was supported in part by the NSFC under Grant 61572449, Grant 61972123, and Grant 61901150, in part by the Key RD projects under Grant 2018YFC0830106, and in part by the ZJNSF under Grant Y19F020124.

ABSTRACT In video coding, rate distortion optimized quantization (RDOQ), a popular version of soft-decision quantization (SDQ), achieves superior coding performance, however is ill-suited for hardware implementation due to its inherent sequential processing. On the other hand, deadzone hard-decision quantization (HDQ) is friendly to hardware implementation, however suffers from non-negligible coding performance degradation. This paper proposes a content-adaptive deadzone offset model to improve the coefficient-wise deadzone HDQ by imitating the behavior patterns of RDOQ. The contributions of this paper are characterized by twofold. On one hand, this work formulates seeking optimal deadzone offset model as a problem of binary classification, and analyzes the distribution characteristics of the optimal deadzone offsets obtained from samples by fully imitating RDOQ, and then derives adaptive deadzone offset model by maximizing the right classification probability of offset-induced rounding in HDQ. On the other hand, the distribution parameters of DCT coefficients are measured in a position-wise way, and the adaptive deadzone model is built by applying Maximum a posterior estimation method according to three characteristic parameters, i.e. quantization step size, parameter of DCT coefficients, and quantization remainder, in the sense of rate distortion optimization. Simulation results verify that the proposed adaptive HDQ algorithm, in comparison with fixed-offset HDQ, achieves 0.54% and 0.52% bit rate saving on average with almost negligible complexity increment. Simultaneously, the proposed algorithm only sacrifices smaller than 0.55% and 0.54% increment in terms of BD-BR in comparison with RDOQ. The proposed HDQ is well-suited for hardwired video coding implementation.

INDEX TERMS Video coding, rate distortion optimization (RDO), soft decision quantization, hard decision quantization.

I. INTRODUCTION

Video coding systems based on H.26x and MPEG standards had been successfully employed in many multimedia applications. In 2013, the JCT-VC (Joint Collaborative Team on Video Coding) had finalized a new video coding standard, called High Efficiency Video Coding (HEVC) [1]. HEVC aims at achieving 50% coding efficiency improvement or more compared to its predecessor H.264/AVC [2]. Like other video codecs, HEVC employs rate-distortion optimization (RDO) [3] to select optimum coding parameters such as prediction modes, motion vectors, transform block sizes, coding block sizes and quantized level of transform coefficients etc. Efficient HEVC codec design suffers from extremely high computation complexity due to complex coding tools used [4].

Quantization plays an important role in determining the rate-distortion (RD) performance of video coding [5], [6]. It not only determines quantization distortion, but also has prominent impacts on coding rate. Since video standards

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

only define the inverse quantization, many research works have explored how to efficiently quantize the DCT (discrete cosine transform) coefficients while remaining compliant with respective to video standards (see for example [7], [8] and references therein).

In early video codecs, DCT coefficients were quantized generally using a uniform scalar quantizer (USQ). Later on, a USQ with deadzone (USQ + DZ) was adopted in MPEG-4 and early H.264/AVC reference codes [7], [9]. In a USQ + DZ, a fixed rounding offset is usually determined empirically in a heuristic way. The quantization process in the case of USQ and USQ + DZ is the so-called hard-decision quantization (HDQ), since given USQ or USQ + DZ, the quantization output in response to an input coefficient is uniquely determined by the input, and correlation among coefficients and their effects on quantization are not considered [7]–[10].

Inspired by [11], soft-decision quantization (SDQ) is a better alternative with superior RD performance contributed by full utilization of correlations among inter-coefficients [8]. A popular SDQ implementation is Viterbi trellis search [8], [10], [12], which was implemented in [12] for H.263+ and in [10] for H.264/AVC. However, running dynamic programming over the full trellis is computationally expensive. To get around this, in H.264/AVC and HEVC reference softwares JM and HM, a simplified SDQ, called rate distortion optimized quantization (RDOQ) [13], was adopted. Almost similar RD performance of full-trellis Viterbi search based SDQ is achieved by RDOQ, in which partial local paths in the trellis graph are searched with considerable computation saving [14].

In comparison with HDQ, SDQ (including RDOQ, similarly hereinafter) is still computationally expensive. In SDQ, optimal quantization result is determined from multiple quantization candidates by comparing their coding costs, i.e. $J = D + \lambda \times R$, in the sense of RD optimization. It is the calculation of RD cost for each candidate that result in the high complexity of SDQ [15]–[18]. Since quantization lies in the innermost loop of RD optimized algorithm, the computation burden of this coding cost calculation is further multiplied by the number of candidate intra and inter coding modes, which are abundant in the latest video standards [4]. Consequently, efficient implementation of SDQ is still an extremely important challenge for video codec optimization.

In the literature, there have been several recent works proposed to alleviate the computation burden of SDQ [17]–[25]. Some scholars proposed all-zero block detection algorithms as an early stage prior to quantization even transform to bypass quantization, reducing the complexity of HEVC video coding [19]–[22]. Other scholars proposed fast quantization algorithm to replace the computation-intensive SDQ algorithm [17], [18], [23]–[25]. For example, focusing on the RD cost difference among adjacent candidates and its impacts, Lee presented a method to reduce the computation of RDOQ using a bypass decision method and a simplified level adjustment method [18]. Other works included reducing the number of candidate quantization results [23], employing fast computation for RD cost of candidate coefficient levels [24], and using fast bit rate evaluation [25]. These methods alleviate the computation burden in SDQ to certain degrees. Nevertheless they are generally designed for soft-targeted video coder optimization. The high complexity and hardware unfriendliness partly stems from complicated rate estimation [26] and sequential processing in Context Adaptive Binary Arithmetic Coding (CABAC) [27]. Some scholars proposed fast rate estimation schemes to simplify the rate estimation [28]–[31]. However, it is not easy to build accurate coefficient level rate model.

It is well known that hardwired throughput is a crucial consideration factor in terms of hardware architecture design, especially for real-time video coder with RD optimization [14], [32]. Parallel processing and pipelining are crucial techniques for efficient implementation with specific throughput constraint on hardware platforms [14], [32]. Efficient pipelining for SDQ is challenged by the inherent data dependencies in trellis search and context state transition in CABAC [14]. The inner data dependency severely hinders SDQ from efficient hardwired implementation. The pipelining efficiency degradation caused by data dependency was not completely considered in traditional SDQ algorithms mentioned above. For example, the hardware constraint caused by data dependency was mentioned in [18] as a motivation, but was not fully addressed. In fact, the method in [18] is still more suitable for software video encoders.

In contrast, there is no data dependency problem in HDQ, such as the prevalent deadzone based HDQ [7], [9]. Coefficient level parallel processing can be achieved in coefficient-wise HDQ via employing hardware pipelining. Unfortunately, there is a nontrivial rate distortion (RD) performance gap between HDQ and SDQ. Huang presented an analytical method to speed up RDO-based quantization by employing a rate model [17]. By ignoring the inter-coefficient dependency, the authors derived a coefficient-independent deadzone model to simplify RDOQ. However, that work is essentially a deadzone HDQ algorithm, as opposed to RDOQ. The momeryless source assumption is inconsistent with the intrinsic dependency in CABAC based SDQ. Moreover, the accuracy of coefficient level rate model in [17] is supposed to be improved further. Otherwise, the RD performance gap as opposed to SDQ is unavoidable. In summary, it is meaningful to further improve the RD performance of HDQ for hardwired video coder, taking the inter-coefficient correlation into account by imitating the behaviors of SDQ.

The authors had made superficial researches on algorithm design for fast SDQ [14] and efficient HDQ [33]. In [14], a fast SDQ algorithm was proposed to decrease the number of trellis search stages to decrease the complexity and to break data dependency in optimal SDQ. The inter-coefficient dependency was considerably decreased in [14], however there are still existing dependency that is harmful for latency-sensitive video coders. In [33], heuristic method statistical analysis was applied to derive relatively accurate deadzone

offset models for improving HDQ algorithm. However, there are still room for performance improvement. This work is an expansion of the previous work in [33].

This paper proposes an optimized deadzone offset model for improving deadzone HDQ. The contributions of this paper are characterized by twofold. On one hand, this work formulates seeking deadzone offset model as a problem of binary classification according to the zero-offset quantization remainder, and applies statistic analysis method to explore the distribution characteristics of the desired deadzone offsets obtained from samples by fully imitating the behavior of RDOQ, and then derives adaptive deadzone offset model by maximizing the right classification probability of offset-induced rounding in HDQ. On the other hand, the distribution parameters of DCT coefficients are measured in a position-wise way, and the adaptive deadzone model is built by applying Maximum a posterior estimation method according to three characteristic parameters in the sense of rate distortion optimization.

The rest of this paper is organized as follows. Problem formulation is given in section II. The proposed offset model and HDQ algorithm is given in section III. Section IV gives the simulation results. Finally, Section V concludes the paper.

II. CHALLENGES AND MOTIVATION

In CABAC-based SDQ, inter-coefficient correlation is explored in conjunction with the subsequent lossless coding. The quantized level of each DCT coefficient depends not only on the coefficient itself, but also on other coefficients in the same block [8], [10], [11], [13], [14]. Dynamic programming is usually employed to implement optimal SDQ, in which a trellis graph is built to describe the inter-coefficient dependency. In the trellis graph, there are multiple stages, one stage for each DCT coefficient in a block, and at each stage there are multiple candidate quantized levels, described as states at that stage. Viterbi search is applied to find an optimal path in the graph with the minimum coding cost. When implemented in this way, SDQ achieves superior RD performance, approximately $6\%-8\%$ rate saving, in comparison with fixed offset deadzone HDQ [10]. The drawbacks of SDQ, however, lie in its high computation complexity and its challenges in hardware implementation [14].

In contrast, DCT coefficients are quantized in an independent, coefficient-by-coefficient manner in HDQ. That is, the quantized level of each coefficient depends only on that coefficient itself. Consequently, HDQ is well-suited for hardware implementation due to friendly support of parallel processing and pipelining. In early video codecs, simple half-rounding HDQ was widely used. Later on, deadzone HDQ with fixed offsets ($\frac{1}{3}$ and $\frac{1}{6}$ for intra and inter modes) were widely used in H.264 JM [34] and HEVC HM codecs [35] due to its matching with entropy coding [7], [36]. The disadvantage of HDQ, however, lies in its inferior RD performance. Therefore, it would be desirable to develop an adaptive deadzone offset model to improve the traditional fixed-offset HDQ by learning the inner characteristics of SDQ [33], [37].

To motivate our proposed model, we now look at deadzone HDQ through the lens of SDQ. In deadzone HDQ, an offset δ is employed to adjust the quantization result z as follows

$$z = floor(\frac{|x|}{q} + \delta) \tag{1}$$

Here, $|x|$ is the intensity level of the DCT coefficient x to be quantized, and q is the quantization step size determined by quantization parameter Qp. In H.264/AVC and HEVC, q is roughly equal to $2^{\frac{(Qp-4)}{6}}$ [38], and floor is the rounding operator. There are four challenges from the systematic perspective of designing optimal deadzone HDQ.

Firstly, the deadzone offset δ in HDQ is supposed to be determined adaptively in a coefficient-wise way in terms of rate distortion optimization, accounting for the context adaptive processing in CABAC. In general, Laplacian distribution is used to model the DCT coefficients due to its good tradeoff between expression concision and model accuracy [5], and the probability density function (PDF) is described by

$$f_L = \frac{1}{2\Lambda} e^{\frac{-|x|}{\Lambda}} \tag{2}$$

where Λ is the parameter of the Laplacian distribution model. Given transformed coefficient sample x_i, the maximum likelihood estimate of Λ can be computed as follows

$$\Lambda = \frac{\sigma}{\sqrt{2}} = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{3}$$

Here, N is the count of the samples collected from multiple frames used for parameter estimation. Based on Laplacian DCT distribution model, the deadzone offset δ of HDQ is typically determined using rate distortion optimization [9].

$$\delta = \frac{q}{2} - \frac{\lambda}{2 \times ln(2) \times \Lambda} \tag{4}$$

Here, λ is the Lagrangian multiplier for rate distortion optimization [38]. Without loss of generality, the normalized deadzone offset is derived as follows.

$$\delta'_L = \frac{\delta_L}{q} = \frac{1}{2} - \frac{\lambda}{2 \times ln(2) \times q \times \Lambda} \tag{5}$$

According to the description above, the accuracy of parameter Λ is crucial for derivation of deadzone offset. Yet, it had been widely proven that coefficients at different positions have different Λ values [36], which means the parameter calculated according to the formula (3) is not accurate enough and thereby coefficient-level quantization manipulation in accordance with coefficient-wise DCT parameter is significant for RD performance improvement of HDQ.

Secondly, the Lagrangian multiplier λ is influential for deadzone offset model as shown in (4). Under the Laplacian model hypothesis [5], [7], λ is derived according to $\lambda = -(\frac{\partial D}{\partial q})/(\frac{\partial R}{\partial q})$ and described as follows [5].

$$\lambda = ln(2)\frac{(q-\delta)^2 - \delta^2}{q} \tag{6}$$

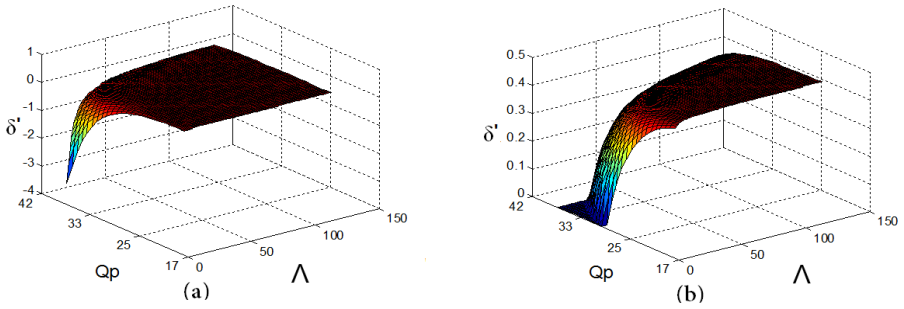

FIGURE 1. (a) deadzone offset δ' (b) deadzone offset δ' (forcing the negative offset to zero).

In fact, the deadzone offset model in (4) is derived from equation (6). The parameters λ and δ are dependent with each other. In view of this dilemma, there is an egg-and-chicken problem here in terms of optimal selection for λ and δ. Moreover, coefficient level solutions can only be theoretically derived from equations (4) and (6), however it is challenging and not easy to be solved. In H.264/AVC and HEVC reference codes [34], [35], JM and HM, the widely used Lagrangian multiplier is expressed as follows [38].

$$\lambda = c \times 2^{\frac{Qp-4}{3}} = c_1 \times q^2 \tag{7}$$

Here, c is a constant experimentally determined as 0.68 or 0.85 [38], and c_1 is a scaling parameter related with c. If applying equation (7) in (5), the normalized deadzone offset can be described as follows.

$$\delta'_L = \frac{1}{2} - \frac{c_1 \times q^2}{2 \times ln(2) \times q \times \Lambda} = \frac{1}{2} - \frac{c_1 \times q}{2 \times ln(2) \times \Lambda} \tag{8}$$

The normalized offset δ'_L is illustrated in Fig.1. In accordance with figure (a), the offset may be even negative sometimes. These special and unnormal cases are caused by the exaggeratedly simplified Lagrangian multiplier shown in (7), which is not derived in an optimal way [39], [40]. We force these negative offsets to zero, and the adjusted offset is given in figure (b). The HDQ algorithm based on the offset in equation (8) suffers from nontrivial RD performance degradation as opposed to anchor SDQ.

Thirdly, the coefficient-wise model in [5] is built in a macroscopic way based on statistic analysis. More importantly, the deadzone offset is estimated for all coefficients without considering the quantization parameter and source model parameter. This fixed constant models, $\frac{1}{3}$ and $\frac{1}{6}$, are used for deadzone offset modeling and incorporated into JM and HM reference codecs. However, SDQ manipulates the quantization result in a microscopic way, specifically by dynamic programming according to the probabilities of the contexts in CABAC. The inter-coefficient influence is not considered in coefficient-level offset model shown in (5). These simplifications in fixed-offset deadzone HDQ will indispensably result in RD performance loss [7], [9], [36]. Yet, the inter-coefficient influence is difficult to analyze, and how to incorporate it into offset modeling is the challenge that should be addressed in this paper.

Fourthly, the quantization remainder of simple division with zero-offset, i.e. the decimal part ξ, has a prominent impact on the quantization distortion and optimal deadzone offset. In this paper, ξ is expressed as follows.

$$\xi = \frac{mod(|x|, q)}{q} \tag{9}$$

Here, *mod* is the remainder operation. The differences of versatile quantization algorithms, such as deadzone HDQ and RDOQ, are equivalently characterized by rounding ξ to different results, e.g. 0 or 1 in brief. The ways of rounding ξ differ in different quantization algorithms. As far as HDQ is concerned, ξ directly influences the deadzone rounding offset. It is indispensable to take ξ into consideration in terms of adaptive offset modeling.

In conclusion, it is meaningful to excavate the inner characteristics of optimal SDQ as guidance to propose adaptive offset model for improving traditional fixed offset HDQ. The goal is to approach the RD performance of SDQ as much as possible and maintain the advantage of parallel processing in HDQ. The challenges encountered in offset modeling are summarized as follows: (1) How to estimate the position-wise Laplacian parameter accurately well-suited; (2) How to resolve the egg-and-chicken problem in the derivation of offset modeling; (3) How to analyze the inter-coefficient influence properly and incorporate it into offset modeling; (4) How to consider multiple parameters, quantization parameter, remainder, DCT distribution parameter jointly with inter-coefficient influence considered in offset modeling.

In order to address above challenges, this paper formulates the offset modeling problem as a binary classification problem according to the quantization remainder for developing accurate offset model. Specifically, the Laplacian parameter is estimated in a position-wise way according to the DCT coefficient distribution. To avoid the egg-and-chicken problem, we model the offset from patterns depicted in massive samples in the statistical way. In HEVC, the probability contexts of the syntax element *level* depend on the numbers of prior coefficients with quantized levels equal to 1 and larger than 1, respectively. So, the inter-coefficient influence is analyzed according to this property and fully utilized in offset modeling. Finally, the content-adaptive offset model is proposed by fusing above mentioned characteristics properly.

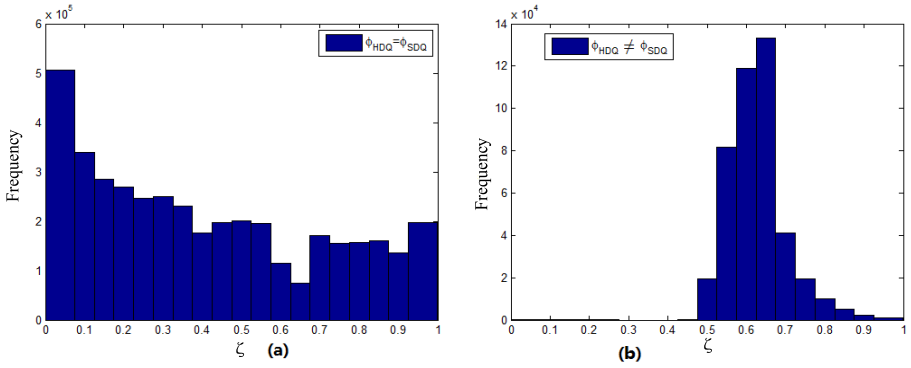

FIGURE 2. The occurring frequency results of $\phi_{HDQ} = \phi_{SDQ}$ (a) and $\phi_{HDQ} \neq \phi_{SDQ}$ (b) in the cases of different ξ.

III. PROPOSED COEFFICIENT-WISE HDQ

A. PROBLEM FORMULATION

In H.265 reference codec HM 16.0, each DCT coefficient is quantized by HDQ or SDQ. We denote $\phi_{HDQ}(\xi)$ and $\phi_{SDQ}(\xi)$ as the rounding results of ξ of HDQ and SDQ, respectively. There are two possible situations according to the condition whether $\phi_{HDQ}(\xi)$ is equal with $\phi_{SDQ}(\xi)$ or not, i.e. $\phi_{HDQ} = \phi_{SDQ}$ and $\phi_{HDQ} \neq \phi_{SDQ}$. Fig.2 gives the results of histogram of these two situations in the case of different ξ. This offline statistic analysis is conducted using massive samples collected from different test sequences and different quantization parameters covering low, median and high bit rate cases. In fact, the final result in RDOQ is obtained by fine-tuning the result of HDQ with fixed offset $\frac{1}{2}$. Specifically, if we apply z_{HDQ} and z_{RDOQ} to represent the results of HDQ and RDOQ, respectively, z_{RDOQ} is equal to z_{HDQ} or $z_{HDQ} - 1$ in most cases, and z_{RDOQ} may be equal to $z_{HDQ} - 2$ when $z_{HDQ} = 2$. According to Fig.2-(b), we observe that the occurring frequencies differ a lot in the case of $\phi_{HDQ} \neq \phi_{SDQ}$, the samples mainly occur in the regions in which ξ ranges from 0.5 to 1.0. This observation verifies that ξ has predominant influence on the rounding results.

Besides, it also can be found that there is a very small probability that $\phi_{HDQ}(\xi) \neq \phi_{SDQ}(\xi)$ occurs when $\xi \in [0, 0.5)$. This case means that samples with ξ in range [0,0.5) may be rounded to negative value -1. Obviously, $\phi_{HDQ}(\xi) \neq \phi_{SDQ}(\xi)$ causes the performance difference between HDQ and SDQ. Yet, the probability of the latter case is smaller than 0.1% as shown in Fig.2. As a result, we mainly focus on the first case, i.e. $\xi \in [0.5, 1)$, and ξ is mainly rounded to 0 or 1 in this case.

In fact, the genuine problem of deadzone HDQ is a classification problem, i.e. rounding the decimal part ξ to -2, -1, 0 or 1. According to the result shown in Fig.2, this problem is simplified to a binary classification problem, i.e., ξ is rounded to 0 or 1. In traditional HDQ, fixed deadzone offset is applied for rounding on ξ. Fig.3 shows the curves of probability density function (PDF), normalized frequency equivalently, in two kinds of samples, $\phi_{SDQ}(\xi) = 0$ and $\phi_{SDQ}(\xi) = 1$, in the case of different remainder ξ and frequency positions. As shown in Fig.3, the PDF curves of ξ differ at frequency positions, and so the adaptive deadzone offsets depend on the frequency position, i.e. DCT distribution parameter. Dynamic offset model is highly desired to be developed.

As analyzed above, it is difficult to derive an analytic solution accounting for the intrinsically complex mechanism behind Viterbi-based SDQ and RDOQ. Instead, this work tends to implement an accurate deadzone offset model according to ξ to achieve successful classification by rounding ξ. This is achieved by comparing SDQ and deadzone HDQ and forcing HDQ to derive the rounding result $floor(\xi + \delta)$ identical with $\phi_{SDQ}(\xi)$ in SDQ. The classification can be expressed as follows.

$$floor(\xi + \delta) = \begin{cases} 0, & \text{if } \phi_{SDQ}(\xi) = 0 \\ 1, & \text{if } \phi_{SDQ}(\xi) = 1 \end{cases} \tag{10}$$

In view of analysis in section II, the optimal deadzone offset will be adaptively determined by statistic modeling taking three crucial parameters into consideration, the coefficient-wise DCT distribution parameter, quantization remainder of fixed offset, and quantization parameter.

Instead of directly adopting the DCT distribution parameter Λ, this work takes the inner-coefficient dependency into account. Λ describes the statistical property of the position-wise coefficients. Here, we instigate the difference between Λ_j and the intensity level of the sample at position with index j, i.e. x_j, and use the difference between these two items to capture the dissimilarity degree between the ensemble and specific samples.

In HEVC, the size of the transform unit (TU) varies from 4×4 to 32×32. The average intensities of DCT coefficients at different positions in different size TUs are shown in Fig.4. According to this figure, we can find that the distribution patterns of the average DCT coefficient intensity in each $4\times$ 4 block in different size TUs are relatively similar. So, we divide each TU into several 4×4 blocks and the Laplacian parameter at j-th position (Λ_j, $j \in [0, 15]$) is calculated according to all samples at j-th position at all TUs. In HEVC, the probability contexts of the intensity levels depends on the numbers of prior quantized coefficients with levels equal to 1

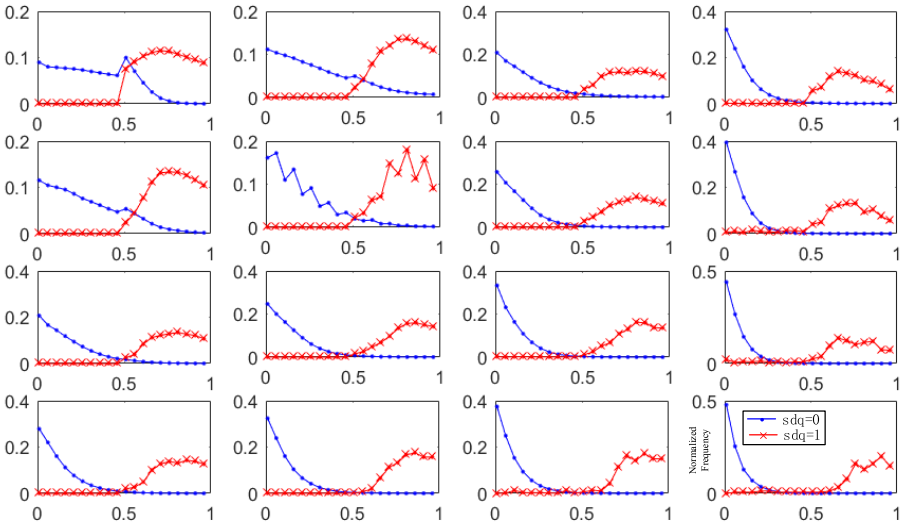

FIGURE 3. The probability density function curves of $\phi_{SDQ} = 0$ and $\phi_{SDQ} = 1$ in the cases of different ξ and frequency positions.

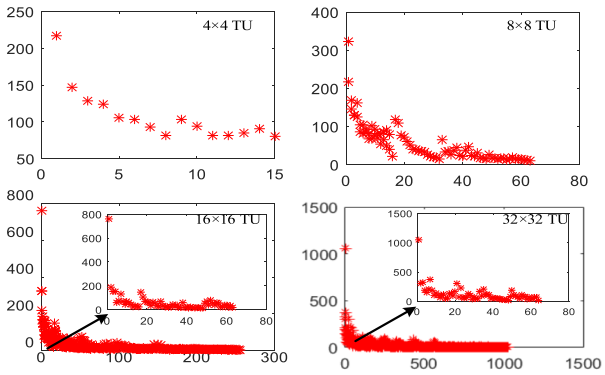

FIGURE 4. DCT coefficient intensities in TUs with size from 4 × 4 to 32 × 32.

and larger than 1 respectively [10], [26]. To coincide with context coding in CABAC, this work defines the accumulated difference η_i by summing up the difference of the remaining coefficients in one block in the ZScan order.

$$\eta_i = \sum_{j=i}^{N_{cof}} ||x_j| - \Lambda_j| \tag{11}$$

Here N_{cof} is the number of the coefficients in one transform block, and $|\cdot|$ is the absolute operator. η_i reflects the difference degree of Λ_i between the ensemble and the coefficients in one block prior to the i-th coefficient. η_i will be employed as opposed to Λ_i for deadzone offset modeling.

To achieve online learning and adaptive generalization, the proposed work updates the parameter Λ by simply averaging, the original parameter Λ_{ori} is updated as the new one Λ_{new} as follows.

$$\Lambda_{new} = \frac{\Lambda_{ori} \times N_{ori} + |x_i|}{N_{ori} + 1} \tag{12}$$

Here, N_{ori} is the count of coefficient samples before the current one. The parameter Λ_{ori} is initialized as the the average value of the previous identical-type frame.

B. STATISTIC ANALYSIS BASED MODEL BUILDING

In view of trellis search based SDQ, it is extremely difficult to achieve 100% right classification for deadzone HDQ successfully by only adjusting deadzone offset δ. The quantization decimal part, i.e. remainder ξ may be rounded to 0 or 1 in SDQ. According to the rounding results, all coefficients can be categorized into two kinds, ''up-rounding'' and ''down-rounding'' coefficients. In the case of ''up-rounding'' and ''down-rounding'', the rounding results of ξ are equal to 1 and 0, i.e. $\phi_{SDQ}(\xi) = 1$ and $\phi_{SDQ}(\xi) = 0$, respectively.

If we apply a deadzone offset model in deadzone HDQ, there are two cases in terms of the rounding results compared with SDQ. On one hand, there are overwhelming majority of the ''up'' and ''down'' rounding coefficients that have identical rounding results as SDQ, and we record them as ''inlier'' samples. On the other hand, there are minor proportions of the ''up'' and ''down'' rounding coefficients that have different rounding results compared with SDQ, and we record them as ''outlier'' samples. It is crucial to minimize the occurring frequencies of ''outliers'' to track the optimality of SDQ. We conducted offline statistic analysis on ''inlier'' and ''outlier'' samples. Fixed offsets $\frac{1}{3}$ and $\frac{1}{6}$ are used for deadzone HDQ in statistic analysis. Fig.5 and Fig.6 give the ξ-wise occurring histogram of inliers and outliers samples in the cases of ''down'' and ''up'' rounding coefficients respectively.

The results in Fig.5 and Fig.6 indicate that the ''inliers'' mainly appear in the range $\xi < 0.6$ and $\xi > 0.7$ in the cases of ''down'' and ''up'' rounding coefficients respectively. In contrast, the ''outliers'' mainly appear in the range $0.5 < \xi < 0.7$ in the case of ''up-rounding'' coefficients caused by the

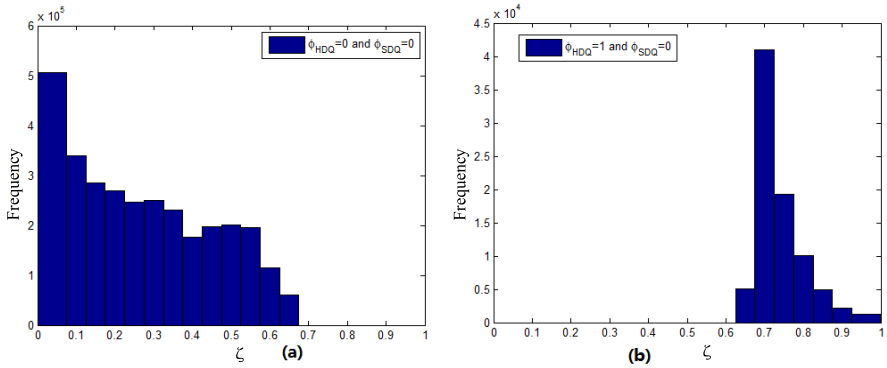

FIGURE 5. The occurring frequency results inliers (a) and outliers (b) of the down-rounding samples (offset $\frac{1}{3}$).

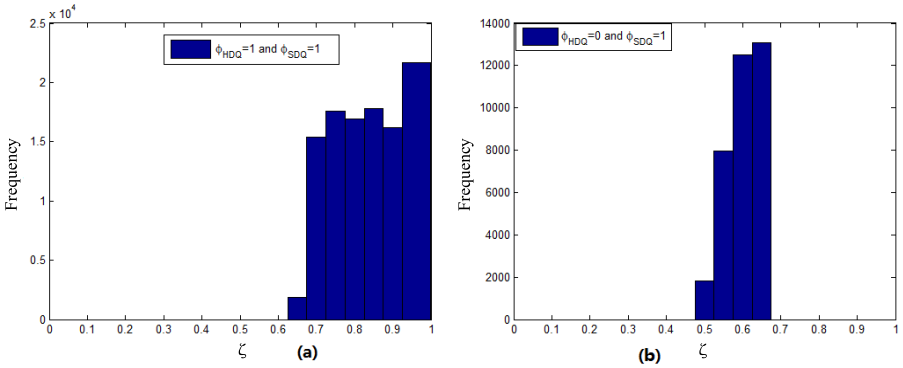

FIGURE 6. The occurring frequency results of inliers (a) and outliers (b) of the up-rounding samples (offset $\frac{1}{3}$).

overregulation due to inappropriately exaggerated deadzone offset. The ''outliers'' mainly appear in the range $0.65 < \xi < 0.9$ in the case of ''down-rounding'' coefficients caused by the under-regulation due to inappropriately shrunken deadzone offset. The occurring frequencies of ''inliers'' determine the accurate classification probability, while the occurring frequencies of the ''outliers'' reflect the wrong classification probability.

Fig.7 shows the relationship between wrong classification probability of ''down-rounding'' samples, p_{wd}, and that of ''up-rounding'', p_{wu}. The rounding results of remainders in SDQ, $\phi_{SDQ}(\xi) = 0$ and $\phi_{SDQ}(\xi) = 1$, are respectively shown in blue and green colors in Fig.7. In contrast, the samples with $\phi_{HDQ}(\xi) = 0$ and $\phi_{HDQ}(\xi) = 1$ are shown in the regions above the dotted line and that below the dotted line respectively. A well-designed offset δ for deadzone HDQ may result in right classification and wrong classification compared with SDQ. Intuitively, this optimal offset model is supposed to be built by maximizing the right classification probability, equivalently minimizing misclassification. We modify HDQ algorithm by seeking an appropriate deadzone offset (δ_{opt}) from all possible candidates (δ_{can}), improving HDQ by deriving identical rounding result as SDQ. This above process can be described as follows.

$$floor(\xi + \delta_{opt}) = \phi_{SDQ}(\xi) \tag{13}$$

It is challenging to seek a well-designed and universal offset δ_{opt} to correctly round all samples as shown in Fig.5 and Fig.6. Aiming at abstracting the common characteristics in the sense of statistic analysis, we investigate the ''outliers'' samples in the cases of ''up-rounding'' and ''down-rounding'' coefficients respectively.

On one hand, the ''outliers'' occur in the case of ''down-rounding'' coefficients as shown in Fig.5-(b). In this case, HDQ fails to track the optimality of SDQ, i.e. $\phi_{HDQ} \neq \phi_{SDQ}$ equivalently. It is observed that the probability of $z_{HDQ} = z_{SDQ} + 1$ is close to 1 when the outlier samples in this case [14] are analyzed. This statistical phenomenon can be justified as follows. In SDQ, decreasing the quantized level by 1 will increase the distortion compared deadzone HDQ with specific offset; however, the increment is well compensated by the resulting bit rate saving, yielding a decreased RD coding cost and superior RD performance. Thus, HDQ improvement can be achieved by seeking an appropriate offset δ_{opt} from candidate offsets δ_{can} for the ''down-rounding'' coefficients described as follows.

$$\delta_{opt} = \underset{\delta_{can}}{arg}\{floor(\xi + \delta_{can}) = 0\} \tag{14}$$

There are many candidate offsets δ_{can} that satisfy the constraint in equation (14), and all possible offsets correspond to a maximal possible offset range (g_{min1}, g_{max1}),

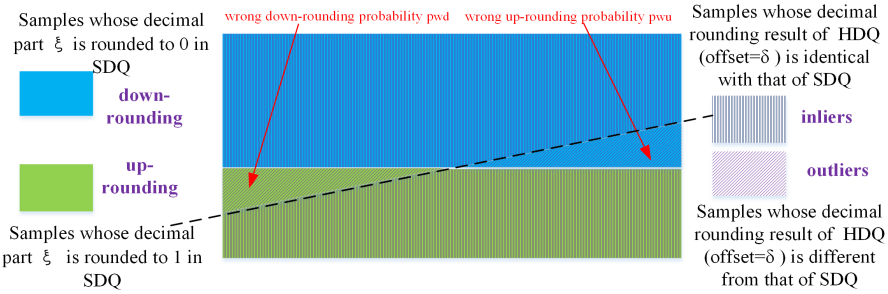

FIGURE 7. The "up-rounding" and "down-rounding" coefficients, "outlier" and "inlier" samples, as well as the wrong down-rounding probability p_{wd} and the wrong up-rounding probability p_{wu} in deadzone HDQ with rounding offset δ.

i.e. down-rounding offset range. The modified HDQ algorithm using a deadzone offset within the range (g_{min1}, g_{max1}) can force the "outliers" of "down-rounding" coefficients to get identical quantized results as SDQ.

On the other hand, the "outliers" may also occur in the case of "up-rounding" coefficients as shown in Fig.6-(b). In this case, HDQ fails to track the optimality of SDQ due to overregulation caused by inappropriate deadzone offset. As for these samples, the offset used in modified HDQ should maintain the similar quantized output of SDQ as much as possible as follows.

$$\delta_{opt} = \underset{\delta_{can}}{arg}\{floor(\xi + \delta_{can}) = 1\} \tag{15}$$

There are multiple candidate offsets δ_{can} that satisfy the condition in equation (15), and all possible offsets correspond to a maximal possible offset range (g_{min2}, g_{max2}), i.e. up-rounding offset range. The modified HDQ using a deadzone offset within range (g_{min2}, g_{max2}) can guarantee the "outliers" of "up-rounding" coefficients to maintain the same quantized results of SDQ, and prevent converting the original "inliers" to "outliers". As a result, the modified HDQ can approach the performance of SDQ as much as possible.

This work employs maximum a posteriori (MAP) method to assist in statistic analysis to estimate proper δ by maximizing the right classification probability. As shown in Fig.8, optimal deadzone offset model is built in a coefficient-wise way by employing MAP based analysis, and independent statistic analysis is individually conducted for the "down-rounding" and "up-rounding" coefficients to derive adaptive offset in the cases different parameter combinations. As analyzed in (5), δ is related with quantization parameter Qp, quantization remainder ξ and coefficient-wise distribution parameter Λ. As a result, the parameters Qp, ξ and Λ of "outliers" and "inliers" are collected, in addition the possible offset ranges (g_{min1}, g_{max1}) and (g_{min2}, g_{max2}) of all samples are also recorded simultaneously, then all these statistical samples will be used for deadzone offset modeling, as shown in Fig.8.

Assume γ_u and γ_d are the classification prior probabilities of "up-rounding" and "down-rounding" samples.

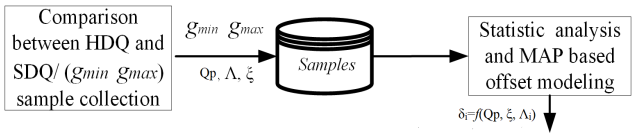

FIGURE 8. Maximum a posteriori estimation (MAP) based derivation of deadzone offset model.

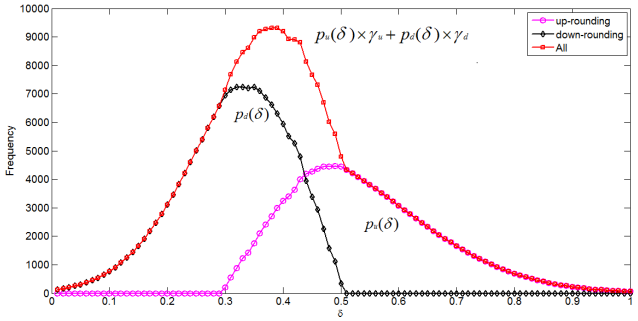

FIGURE 9. The illustration of offset selection under the constraint of maximizing the right classification probability.

Given offset δ, we record $p_{wu}(\delta)$ and $p_{wd}(\delta)$ as the misclassification probabilities of the "up" and "down" rounding samples respectively, and record $p_u(\delta)$ and $p_d(\delta)$ as the conditional probabilities of right classification of the "up" and "down" rounding samples respectively. Apparently, these four items satisfy the following relationship: $p_u(\delta) = 1 - p_{wu}(\delta)$ and $p_d(\delta) = 1 - p_{wd}(\delta)$. According to the conditional probability estimated, seeking an appropriate offset from all candidates can be achieved using MAP as illustrated in Fig.9 via statistic analysis, i.e. maximizing the right classification posterior probability described as follows.

$$\delta_{opt} = \underset{\delta_{can}}{argmax}\{p_u(\delta_{can}) \times \gamma_u + p_d(\delta_{can}) \times \gamma_d\} \tag{16}$$

C. STATISTICAL ANALYSIS OF OPTIMAL OFFSET

As analyzed above, maximizing the right classification probability is in connection with the "outliers" in the case of "down" and "up" coefficients. The resulting offsets are derived by adjusting the deadzone offsets to change the quantization results of deadzone HDQ for these two kinds of

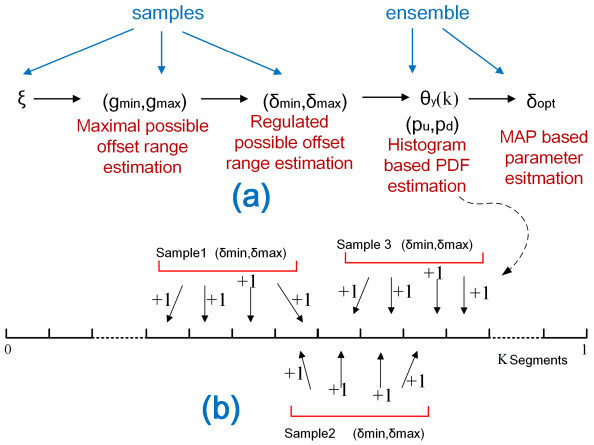

FIGURE 10. The procedures of optimal offset derivation.

"outliers", and a maximal possible offset range (g_{min1}, g_{max1}) and (g_{min2}, g_{max2}) are estimated respectively as follows.

$$(g_{miny}, g_{maxy}) = \underset{\delta_{cany}}{arg}\{floor(\xi + \delta_{cany}) = \phi_{SDQ}(\xi)\} \tag{17}$$

Here, $y = 1$ and $y = 2$ indicate the "down" and "up" rounding samples respectively. For example, (g_{miny}, g_{maxy}) will be (0, 0.45) when ξ is assigned to 0.55 with down rounding. As shown in equation (17), the maximal possible offset range (g_{miny}, g_{maxy}) is obtained by forcing $floor(\xi + \delta_{cany})$ to get the identical rounding results as SDQ, i.e. $\phi_{SDQ}(\xi)$, as shown in equation (14) and (15).

In order to avoid possible offset range overestimation, we employ the regulated possible offset range for deriving offset model under a certain predefined error classification probability allowed. As shown in Fig.10-(a), the maximal possible offset range (g_{min1}, g_{max1}) and (g_{min2}, g_{max2}) are firstly mapped to the so-called regulated possible offset range $(\delta_{min1}, \delta_{max1})$ and $(\delta_{min2}, \delta_{max2})$. Suppose p_{err1} and p_{err2} are the predefined peak values of wrong classification probabilities p_{wd} and p_{wu} allowed for failing to track the optimality of SDQ, i.e. the probabilities of $1 - p_{err1}$ and $1 - p_{err2}$ are achieved in terms of correcting the "outliers" of two cases. According to the predefined wrong probability p_{err1} and p_{err2}, the regulated offset range $(\delta_{miny}, \delta_{maxy})$, i.e. $(\delta_{min1}, \delta_{max1})$ and $(\delta_{min2}, \delta_{max2})$ are determined from the ensemble of (g_{miny}, g_{maxy}) sample, denoted as $(\boldsymbol{g_{miny}}, \boldsymbol{g_{maxy}})$, using statistic analysis as follows.

$$\begin{aligned}(\delta_{miny}, \delta_{maxy}) &= \underset{\delta \in (\boldsymbol{g_{miny}}, \boldsymbol{g_{maxy}})}{arg} \{\widehat{p} = 1 - p_{erry}\} \\ \widehat{p} &= prob\{\phi_{HDQ}(\xi + \delta) = \phi_{SDQ}(\xi)\}\end{aligned} \tag{18}$$

Then, it is crucial to estimate the prior probability density functions (PDF) of the right correction and right maintaining, $p_u(\delta)$ and $p_d(\delta)$, in the case of different δ. Statistic analysis is carried out using a great amount of "down" and "up" rounding samples and their offset ranges $(\delta_{min1}, \delta_{max1})$ and $(\delta_{min2}, \delta_{max2})$. It should be mentioned that the training sequences used have a certain degree influence on the statistic analysis results. In order to guarantee the adaptivity of the proposed work, we employ a variety of test video sequences for offline statistic analysis. In addition, different target rate cases are observed by adjusting the quantization parameters ranging from 18 to 38 with stride 4. This simulation configuration is also used for all other test cases in this work.

Algorithm 1 Brute-Force Search for PDF Estimation

Input: $[\delta_{min1}(m_1), \delta_{max1}(m_1)]$, $[\delta_{min2}(m_2), \delta_{max2}(m_2)]$;
Output: $\theta_1(k)$, $\theta_2(k)$;

Begin
for m_1=1:1:M_1 **do**
  for k=1:1:K **do**
    Begin the Iteration for Intervals
    if $(\frac{(k-1)}{K}, \frac{k}{K}) \in (\delta_{min1}(m_1), \delta_{max1}(m_1))$ **then**
      $\theta_1(k) \longleftarrow \theta_1(k) + 1$;
    else
      $\theta_1(k) \longleftarrow \theta_1(k)$;
for m_2=1:1:M_2 **do**
  for k=1:1:K **do**
    Begin the Iteration for Intervals
    if $(\frac{(k-1)}{K}, \frac{k}{K}) \in (\delta_{min2}(m_2), \delta_{max2}(m_2))$ **then**
      $\theta_2(k) \longleftarrow \theta_2(k) + 1$;
    else
      $\theta_2(k) \longleftarrow \theta_2(k)$;
return $\theta_1(k)$, $\theta_2(k)$;

In this paper, histogram based non-parametric analysis is employed to estimate the PDF curves of $p_u(\delta)$ and $p_d(\delta)$. As shown in Fig.10-(b), the globally maximal possible solution range of $(\delta_{min}, \delta_{max})$, i.e. [0,1], is partitioned into K intervals indexed by k, for example $K = 256$ is used. Then, the possible ranges $(\delta_{min}, \delta_{max})$ of all samples are compared for grouping and classification respectively. As shown in Fig.9-(b), if the regulated offset range falls into one interval indexed by k, the probability of right classification, i.e. the histogram count $\theta_y(k)$, is increased by 1. For the collected samples, $y = 1$ and $y = 2$ correspond to "down-rounding" and "up-rounding" samples respectively. This histogram statistic analysis is described as follows.

$$\theta_y(k) = \begin{cases} \theta_y(k)+1, & (\frac{(k-1)}{K}, \frac{k}{K}) \in (\delta_{miny}(m_y), \delta_{maxy}(m_y)) \\ \theta_y(k), & otherwise(y = 0, 1) \end{cases} \tag{19}$$

Algorithm 1 gives the algorithm flow of brute-force search for PDF estimation. Suppose there are M_1 and M_2 "down" and "up" rounding samples indexed by m_1 and m_2, $[\delta_{min1}(m_1), \delta_{max1}(m_1)]$, $[\delta_{min2}(m_2), \delta_{max2}(m_2)]$ are input to the PDF estimation algorithm, and two occurring frequency results $\theta_1(k)$ and $\theta_2(k)$ are estimated using simple brute-force search method.

The interval-wise histogram results of candidate deadzone offset range are obtained individually for "down-rounding" and "up-rounding" coefficients in the case of different

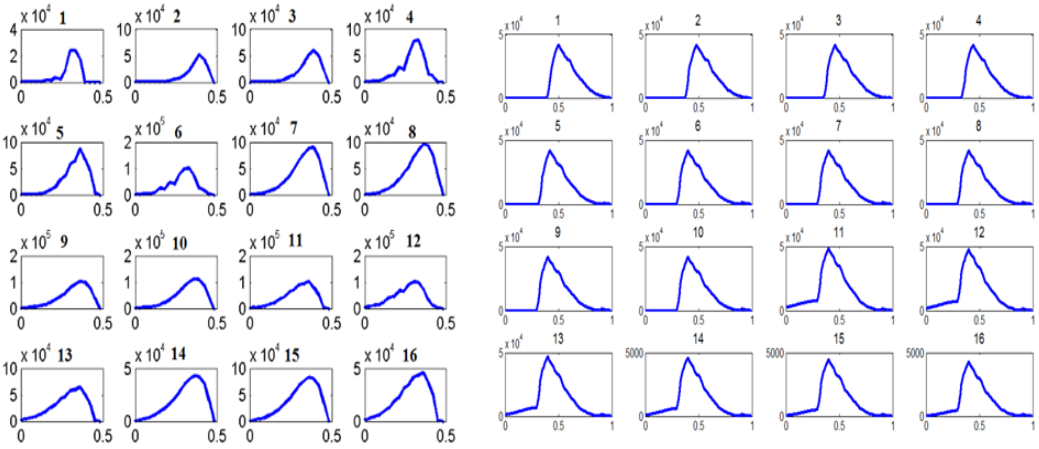

FIGURE 11. (a) The PDF $\theta_1(k)$ of "down-rounding" samples. (b). The PDF $\theta_2(k)$ of "up-rounding" samples.

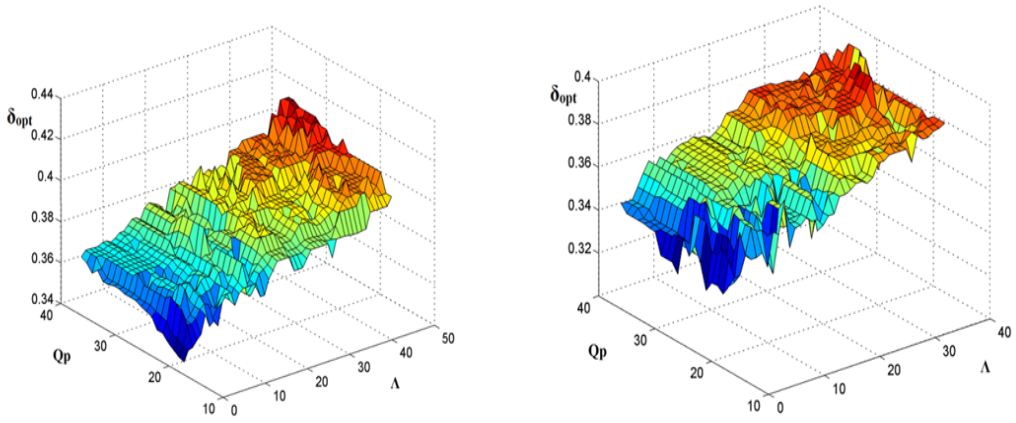

FIGURE 12. (a) δ_{opt} of Intra blocks in the case of $\xi \in (0.5, 0.6)$ (b) δ_{opt} of Inter blocks in the case of $\xi \in (0.5, 0.6)$.

combinations of parameters Qp, ξ, and frequency wise Λ_i. Moreover, statistic analyses are conducted independently in intra and inter coding modes. Typical results of $\theta_1(k)$ and $\theta_2(k)$ of different coefficients are shown in Fig.11-(a) and (b) respectively. Then, $p_u(\delta)$ and $p_d(\delta)$ can be estimated according to $\theta_1(k)$ and $\theta_2(k)$ described in the next subsection.

D. THE ADAPTIVE DEADZONE OFFSET MODEL

In accordance with the characteristics of statistic analysis, this paper builds an adaptive model for rounding offset δ using the MAP method. The proposed model is derived as a function of the combined parameter $\vec{\boldsymbol{p}}$, including quantization parameter Qp, the quantization remainder ξ and the DCT distribution parameter Λ, i.e.

$$\delta = f(\vec{\boldsymbol{p}}) = f(\underbrace{Qp, \xi, \Lambda}_{\vec{\boldsymbol{p}}}) \tag{20}$$

As analyzed above, $\theta_1(k)$ and $\theta_2(k)$ actually reflect the right classification probabilities of maintaining the SDQ quantized results of the "down" and "up" rounding samples, when δ fall into the range $[\frac{(k-1)}{K}, \frac{k}{K}]$. Here, we apply normalization in the histogram results to derive the normalized probabilities, $\theta_1'(\vec{\boldsymbol{p}})$ and $\theta_2'(\vec{\boldsymbol{p}})$, which both are functions of Qp, ξ and Λ as shown in equation (20). The normalized $\theta_1'(\vec{\boldsymbol{p}})$ and $\theta_2'(\vec{\boldsymbol{p}})$ are just the right classification probability of the "down" and "up" rounding samples, i.e. p_d and p_u respectively, in the cases of different parameter combinations. It is well-known that Λ is related with the coefficient position index i, i.e. Λ_i. Under the principle of maximizing the right classification probabilities guided by MAP, we need to seek appropriate offsets to minimize the wrong classification probabilities, i.e. forcing HDQ with the proposed offset model to get the identical results with SDQ as most as possible. Therefore, the peak of histogram probabilities is exploited to determine the optimal deadzone offset shown in equation (21).

$$\delta_{opt}(\underbrace{Qp, \xi, \Lambda}_{\vec{\boldsymbol{p}}}) = \underset{\delta_k(\vec{\boldsymbol{p}})}{argmax}\{\underbrace{\theta_1'(k, \vec{\boldsymbol{p}})}_{p_d} \times \gamma_d + \underbrace{\theta_2'(k, \vec{\boldsymbol{p}})}_{p_u} \times \gamma_u\} \tag{21}$$

The surf results of parameter δ_{opt} in the case of $\xi \in (0.5, 0.6)$, of intra and inter prediction modes, in the case

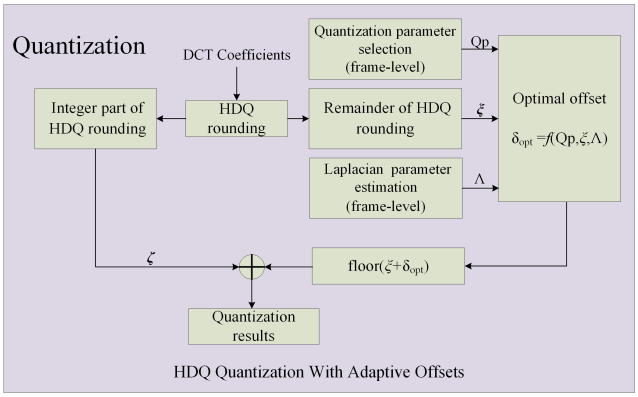

FIGURE 13. The flowchart of the proposed algorithm.

of different combinations between Qp and Λ, are shown as illustration examples in Fig.12-(a) and (b) respectively. The adaptive tables of intra and inter modes are respectively applied in video coding.

The proposed deadzone offset model is combined with equation (1) and finally applied in adaptive deadzone offset HDQ to replace the fixed-offset HDQ. The adaptive models are built via statistic analysis, and the improved HDQ algorithm is still processed in a coefficient-wise way. Thus, HDQ with the proposed offset model still maintains the advantage of friendliness of hardware implementation in HDQ. The models are estimated offline in discrete grid style, i.e. discretizing Qp varying from 12 to 42 with stride 2 (11 discrete intervals), discretizing ξ varying from 0.05 to 0.95 with stride 0.1 (10 discrete intervals), and discretizing Λ varying from 5 to 45 with stride 3 (14 discrete intervals). As a result, we build an offset table with size of $11 \times 10 \times 14 = 1540$, the estimated offsets are expressed in the forms of integer with 8-bit precision scaled from float value. Thus, a buffer with size of 2048 bytes are desired to buffer the whole offset table. The offsets of other characteristic parameter combinations misaligned from the grid in table are linearly interpolated by averaging two adjacent points stored in the table.

This work also support online offset model refinement. The tabularized offsets estimated offline are used as initials. Then, the first GOP can be coded with RDOQ to implement online model training. The statistical analysis method described above is implemented on the first GOP online. The initials will be used to fasten the search of PDF estimation in Algorithm 1, by employing local refinement centered about the initials to avoid brute-force search mentioned in Algorithm 1.

E. OVERALL ALGORITHM FLOW

The flowchart of our algorithm in shown in Fig.13. As shown in this figure, DCT coefficients are quantized by HDQ at first. The quantization integer (ζ) and remainder(ξ) of HDQ rounding are obtained. The optimal offset is calculated according to the equation (21) with three parameters. And finally, the quantization results are ontained by adding the integer and $floor(\xi + \delta_{opt})$. Therefore, compared with the fixed-offset HDQ algorithm, the additional computation comes from two part, parameter Λ estimation and δ_{opt} calculation. Parameter Λ is estimated one time for one frame, online update only consume simple calculation. In addition, regardless of the initial online training, δ_{opt} are built offline and its calculation only needs simple function call as shown in equation (21) or tabulation shown in Fig.12, thereby it is moderate and trivial. As a result, the proposed algorithm is well-suited for parallel hardware implementation.

IV. EXPERIMENTAL RESULTS

A. PERFORMANCE AND ACCURACY

The proposed deadzone offset model and the improved deadzone HDQ are verified in H.265 reference codec HM 16.0. The simplified RDOQ proposed by Huang and Chen [17], the simplified version SDQ i.e. RDOQ [13], and the conventional deadzone HDQ [34] are all taken as performance comparison anchors. Due to this paper is an expansion of our previous work in [33], the corresponding deadzone HDQ proposed in [33] is also as performance comparison anchor. These quantization algorithms are applied in both the final mode coding and the rate distortion optimized mode decision loop. In the training, seven sequences from D1,720p and 1080p format are applied to train our adaptive offset model. And the test sequences for RD performance evaluation include the whole range of HEVC standard test sequences in common test conditions (CTC) [41]. Rate control is turned off, and the quantization parameters 22, 27, 32 and 37 are used for simulation. The PSNR degradation (BD-PSNR) and rate increment percentage (BD-RATE) are used for RD performance comparison [42], [43]. Table 1 and Table 2 give the detailed BD-PSNR and BD-RATE results [42] of all cases in training sequences and test sequences.

In Table 1 and Table 2, RDOQ is taken as the comparison anchor for BD-RATE and BD-PSNR evaluation. As shown in Table 1, compared with RDOQ, the proposed algorithm only has smaller than 0.83% (AI: all I) and 0.67% (RA: random access) BD-BR increment on average in the training sequences. And in the test sequences, BD-BR of proposed algorithm in two cases is 0.55% and 0.54% respectively. Compared with the conventional deadzone, the coding gain of proposed deadzone is $-0.83\% = 0.83\% - 1.66\%$ and $-0.70\% = 0.67\% - 1.37\%$ in training sequences and that is $-0.54\% = 0.55\% - 1.09\%$ and $-0.52\% = 0.54\% - 1.06\%$ in test sequences. Obviously, compared with the test sequences, the training sequences can achieve higher coding gain.

Intensive results also show that in RA case the proposed algorithm only has smaller than and 0.54% BD-BR increment on average in comparison with RDOQ in test dataset, whereas the average BD-BR, in comparison with RDOQ, of the conventional deadzone and deadzone with the model in [17] are 1.06% and 0.84% respectively. Relatively, larger RD performance improvements are observed in intra coding

TABLE 1. The performance comparison of the deadzone HDQ algorithms, offset of Huang [17], the proposed adaptive offset and offset in our previous work [33] compared with optimal RDOQ in training sequences.

Algorithm	Conventional Deadzone				Huang's Deadzone				Wang's Deadzone				Proposed Deadzone			
Sequence	BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)	
	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA
ICE	-0.050	-0.043	0.61	0.57	-0.045	-0.036	0.43	0.45	-0.038	-0.032	0.33	0.32	-0.033	-0.026	0.31	0.28
Cyclists	-0.054	-0.038	0.65	0.68	-0.048	-0.032	0.46	0.54	-0.041	-0.028	0.36	0.38	-0.035	-0.023	0.33	0.33
ElideEditing	-0.111	-0.122	0.88	0.91	-0.098	-0.102	0.62	0.72	-0.084	-0.090	0.48	0.51	-0.072	-0.073	0.44	0.44
Optis	-0.046	-0.046	0.47	0.77	-0.041	0.038	0.33	0.61	-0.035	-0.017	0.26	0.43	-0.030	-0.014	0.24	0.37
Kimono1	-0.117	-0.091	4.27	3.05	-0.103	-0.076	3.00	2.41	-0.089	-0.067	2.33	1.71	-0.076	-0.055	2.15	1.48
ParkScene	-0.032	-0.017	0.92	0.57	-0.028	-0.014	0.65	0.45	-0.024	-0.012	0.50	0.32	-0.021	-0.010	0.47	0.28
pedestrian_area	-0.131	-0.100	3.8	3.03	-0.115	-0.084	2.67	2.40	-0.099	-0.074	2.08	1.70	-0.085	-0.060	1.91	1.47
TOTAL AVE.	-0.077	-0.065	1.66	1.37	-0.068	-0.044	1.17	1.08	-0.059	-0.046	0.91	0.77	-0.050	-0.037	0.83	0.67

TABLE 2. The performance comparison of the deadzone HDQ algorithms, offset of Huang [17], the proposed adaptive offset and offset in our previous work [33] compared with optimal RDOQ in test sequences.

Algorithm	Conventional Deadzone				Huang's Deadzone				Wang's Deadzone				Proposed Deadzone			
Sequence	BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)		BD-PSNR (dB)		BD-BR (%)	
	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA	AI	RA
Traffic	-0.032	-0.019	0.73	0.64	-0.028	-0.016	0.51	0.51	-0.024	-0.014	0.40	0.36	-0.021	-0.011	0.36	0.31
PeopleOnStreet	-0.025	-0.008	0.54	0.19	-0.022	-0.007	0.38	0.15	-0.019	-0.006	0.30	0.11	-0.016	-0.005	0.27	0.09
Kimono1	-0.117	-0.091	4.27	3.05	-0.103	-0.076	3.00	2.41	-0.089	-0.067	2.33	1.71	-0.076	-0.055	2.15	1.48
ParkScene	-0.032	-0.017	0.92	0.57	-0.028	-0.014	0.65	0.45	-0.024	-0.012	0.50	0.32	-0.021	-0.010	0.47	0.28
Cautus	-0.027	-0.031	0.88	1.42	-0.024	-0.026	0.62	1.12	-0.021	-0.023	0.48	0.79	-0.018	-0.019	0.44	0.69
BasketballDrive	-0.046	-0.061	2.36	3.17	-0.041	-0.051	1.66	2.51	-0.035	-0.045	1.29	1.78	-0.030	-0.037	1.19	1.54
BQTerrace	-0.026	-0.019	0.54	1.02	-0.023	-0.016	0.38	0.81	-0.020	-0.014	0.30	0.57	-0.017	-0.011	0.27	0.50
BasketballDrill	-0.010	-0.030	0.16	0.81	-0.009	-0.025	0.11	0.64	-0.008	-0.022	0.09	0.45	-0.007	-0.018	0.08	0.39
BQMall	-0.022	-0.030	0.48	0.75	-0.019	-0.025	0.34	0.59	-0.016	-0.022	0.26	0.42	-0.014	-0.018	0.24	0.36
PartyScene	-0.102	-0.010	2.03	0.32	-0.090	-0.008	1.43	0.25	-0.077	-0.007	1.11	0.18	-0.067	-0.006	1.02	0.15
RaceHorsesC	-0.020	-0.008	0.41	0.24	-0.018	-0.007	0.29	0.19	-0.015	-0.006	0.23	0.13	-0.013	-0.005	0.21	0.12
BasketballPass	-0.034	-0.020	0.74	0.46	-0.030	-0.017	0.52	0.36	-0.026	-0.015	0.40	0.26	-0.022	-0.012	0.37	0.22
BQSquare	-0.006	-0.078	0.09	1.85	-0.005	-0.065	0.06	1.46	-0.004	-0.057	0.05	1.04	-0.004	-0.047	0.04	0.90
BlowingBubbles	-0.051	-0.008	1.19	0.56	-0.045	-0.007	0.84	0.44	-0.039	-0.006	0.65	0.31	-0.033	-0.005	0.60	0.27
RaceHorses	-0.018	-0.008	0.34	0.29	-0.016	-0.007	0.24	0.23	-0.014	-0.006	0.19	0.16	-0.012	-0.005	0.17	0.14
FourPeople	-0.022	-0.049	0.47	1.35	-0.019	-0.041	0.33	1.07	-0.016	-0.036	0.26	0.76	-0.014	-0.029	0.24	0.66
Johnny	-0.056	-0.029	1.68	1.21	-0.049	-0.024	1.18	0.96	-0.042	-0.021	0.92	0.68	-0.036	-0.017	0.84	0.59
KristenAndSara	-0.039	-0.038	0.92	1.10	-0.034	-0.032	0.65	0.87	-0.029	-0.028	0.50	0.62	-0.025	-0.023	0.47	0.54
BasketballDrillText	-0.090	-0.042	2.22	1.02	-0.079	-0.035	1.56	0.81	-0.068	-0.031	1.21	0.57	-0.058	-0.025	1.12	0.50
SlideEditing	-0.111	-0.122	0.88	0.91	-0.098	-0.102	0.62	0.72	-0.084	-0.090	0.48	0.51	-0.072	-0.073	0.44	0.44
SlideShow	-0.060	-0.115	0.95	1.34	-0.053	-0.096	0.67	1.06	-0.046	-0.085	0.52	0.75	-0.039	-0.069	0.48	0.65
TOTAL AVE.	-0.045	-0.040	1.09	1.06	-0.040	-0.033	0.76	0.84	-0.034	-0.029	0.59	0.60	-0.029	-0.024	0.55	0.54

mode compared with inter mode. That is to say that the RD performance loss compared with RDOQ is relatively more contributed by the intra coding blocks. In AI case, more coefficients are quantized to nonzero, so the deadzone offset rounding efficiency degrade a little. As a result, the RD performance results of the AI case are relatively higher than those of the RA case. As shown in Table 2, in AI case, the proposed algorithm only has smaller than 0.55% BD-BR increment on average in comparison with RDOQ, whereas the average BD-BR, in comparison with RDOQ, of the conventional deadzone and deadzone with the model in [17] are 1.09% and 0.76% respectively. In comparison with the conventional deadzone HDQ algorithm, the proposed algorithm achieves 0.54% and 0.52% decrease in terms of BD-BR on average in the cases of AI and RA format respectively. In addition, it is worth noting that the proposed deadzone HDQ in this paper has $0.04\% = 0.59\% - 0.55\%$ and $0.06\% = 0.60\% - 0.54\%$ BD-BR smaller on average than the algorithm proposed in our previous work [33]. The main reason for performance gain is that we take the inter-coefficient dependency into count and update the DCT distribution parameter from Λ_{ori} to Λ_{new} in offset modeling and coding processes. Besides, different from [33] using simple MAP method to estimate δ, this paper employs the regulated possible offset range for deriving offset model under a certain predefined error classification probability. In summary, HDQ based on the proposed adaptive offset is obviously superior to the conventional HDQ, and has relatively superior RD performance compared with the deadzone model in [17], and has only small RD performance degradation compared with RDOQ. The rate distortion curves of the 'Kimono1' sequence are taken as example shown in Fig.14.

The right and wrong classification probabilities can also be used to evaluate the effectiveness of candidate HDQ algorithms. The detailed probabilities of right and wrong classifications relative to RDOQ are given in table 3. As shown in this Table, we choose six sequences in 1080p format as the example. According to the results in table 3, in comparison with fixed-offset HDQ and conventional deadzone HDQ with offset in Fig.1, HDQ with the proposed deadzone offset model can achieve relatively higher right classification probabilities, including right correction and right maintenance, $89.4\% - 84.6\% = 4.8\%$ and $89.4\% - 85.8 = 3.6\%$ increment respectively, simultaneously achieve relatively lower wrong classification probabilities, including wrong correction and maintenance, 4.8% and 3.6% wrong probability reduction

TABLE 3. The right and wrong classification probabilities of candidate algorithms compared with RDOQ.

Test Sources		Fixed offset		Adaptive offset in Fig.1		$\Delta_{P1}(\%)$	Proposed adaptive offset		$\Delta_{P2}(\%)$	$\Delta_{P2}-\Delta_{P1}$
Type	Sequence	Wrong(%)	Right(%)	Wrong(%)	Right(%)		Wrong(%)	Right(%)		
Training Dataset	Kimono1	16.6	83.4	15.2	84.8	1.4	12.4	87.6	4.2	2.8
	Parkscene	13.0	86.9	12.2	87.8	0.8	7.4	92.6	5.6	4.4
	pedestrian_ area	17.4	82.6	16.1	82.9	1.3	12.8	87.2	4.6	3.3
Test Dataset	Cautus	14.9	85.1	12.8	87.2	2.1	9.5	90.5	5.4	3.3
	BacketballDrive	14.9	85.1	13.6	86.4	1.3	9.9	90.1	5	3.7
	BQTerrace	15.8	84.2	14.6	85.4	1.2	11.6	88.4	4.2	3.0
AVERAGE		15.4	84.6	14.1	85.8	1.4	10.6	89.4	4.8	3.4

TABLE 4. The right and wrong classification probabilities of the proposed algorithm and deadzone HDQ.

Sample Type		Classification probabilities of outlier samples				Classification probabilities of Inlier samples			
Test Sources		Conventional Deadzone		Proposed Deadzone		Conventional Deadzone		Proposed Deadzone	
Type	Sequence	Fail to Correct(%)	Success to Correct(%)	Fail to Correct(%)	Success to Correct(%)	Fail to Maintain(%)	Success to Maintain(%)	Fail to Maintain(%)	Success to Maintain(%)
Training Dataset	Kimono1	22.0	78.0	41.0	59.0	12.9	87.1	8.9	91.1
	Parkscene	15.9	84.1	33.5	66.5	10.5	89.5	5.0	94.3
	pedestrian_ area	22.8	77.2	39.5	60.5	13.3	86.7	9.2	90.8
Test Dataset	Cautus	20.5	79.5	38.9	61.1	10.4	89.6	6.6	93.4
	BacketballDrive	20.1	79.9	38.7	61.3	11.5	88.5	7.2	92.8
	BQTerrace	20.9	79.0	39.9	60.2	10.7	89.3	6.7	93.3
AVERAGE		20.4	79.6	38.6	61.4	11.6	88.5	7.3	92.6

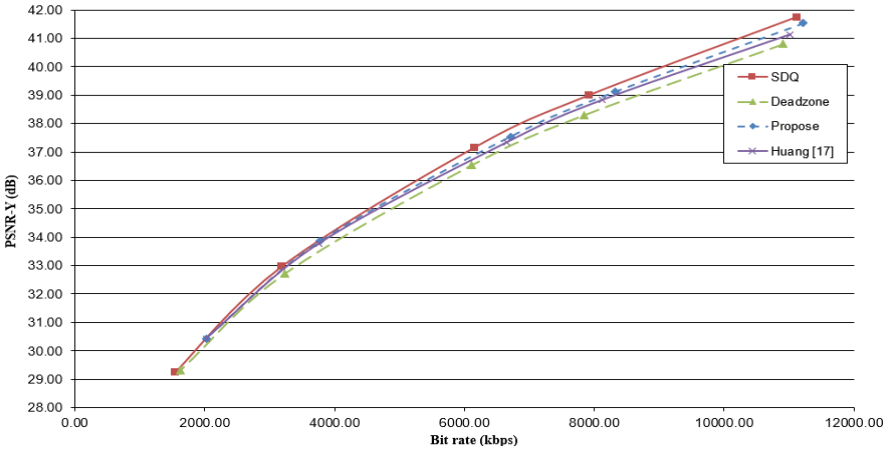

FIGURE 14. RD performance comparisons of the proposed deadzone HDQ and three candidate comparison anchors.

respectively. Extensive simulations are conducted in the case of different quantization parameters, varying from 18 to 38 with stride 4. In order to show the performance of the proposed algorithm clearly, we calculate the right classification probabilities increment, i.e., Δ_P. As shown in Table 3, Δ_{P_1} and Δ_{P_2} indicate the right classification probabilities increments of offset in Fig.1 and the proposed offset compared to the fixed offset. It can be found that Δ_{P_2} falls in the range of 4% to 6%. $\Delta_{P2}-\Delta_{P1}$ is the right classification probabilities increment that results from our adaptive offset. The average value of $\Delta_{P2}-\Delta_{P1}$ is 3.5% and 3.3% in training sequences and test sequences respectively, which is consistent with results in Table 1 and Table 2.

More specifically, the classification probabilities in the case of ''outliers'' and ''inliers'' are given in table 4. These detailed results offer insights into the effectiveness of candidate algorithms. The left half columns of this table illustrate the probabilities of failing to correct and succeeding to correct the results of the ''outlier'' samples compared with RDOQ. The right half columns of this table illustrate the probabilities of failing to maintain and succeeding to maintain the results of the ''outlier'' samples compared with RDOQ. The higher probabilities of ''success to correct'' and ''success to maintain'' indicate better coding performances.

According to the results in table 5, in comparison with fixed-offset HDQ and deadzone HDQ with offset in Fig.1, HDQ with the proposed offset model can achieve relatively high right classification probabilities, and simultaneously achieve relatively low misclassification probabilities, in a way similar with table 3. Moreover, the proposed model based HDQ have a tendency to maintain the ''inlier'' samples as much as possible, and the probability of right maintaining for ''inliers'' is nearly 92.6%. Certainly, this improvement is achieved at the cost of relatively low improvement

TABLE 5. The average running time consumed by one frame coding for all I and I/P/B formats (Unit: min).

Class	AI						RA					
	Deadzone	RDOQ	Huang	Proposed	ΔT_H (%)	ΔT_P (%)	Deadzone	RDOQ	Huang	Proposed	ΔT_H (%)	ΔT_P (%)
A	1.24	8.10	1.59	1.46	18	28	0.80	7.55	1.01	0.93	16	26
B	0.61	3.94	0.77	0.74	22	25	0.40	3.74	0.48	0.46	14	18
C	0.12	0.84	0.16	0.16	29	32	0.11	0.67	0.13	0.14	24	30
D	0.03	0.21	0.04	0.04	33	33	0.28	0.17	0.33	0.34	18	20
E	0.19	1.22	0.25	0.24	26	28	0.17	1.10	0.20	0.19	12	20
F	0.21	1.26	0.26	0.25	22	23	0.18	1.12	0.21	0.20	9	15

regarding the probability of right correcting the ''outliers''. It is reasonable that it is challenging to increase these two probabilities simultaneously in general. The rate distortion performance improvement of the proposed model based HDQ is contributed by maintaining the right results of ''inliers'' and correcting the wrong results of ''outliers'' as much as possible simultaneously. The deadzone HDQ with offset in Fig.1 has a tendency of employing an excessive offset to forcing more coefficients to be quantized to smaller level. However, this adaptive rounding is a little bit exaggerated. In comparison, the proposed algorithm achieves better trade off in terms of correctly rounding on the ''inliers'' and ''outliers''.

B. ALGORITHM COMPLEXITY

As for complexity, Table 5 gives the average running time of quantization in the cases of two GOP structure configurations, AI frames and RA frames, consumed by one frame coding of different algorithms. RDO loop is taken into count for complexity summarization. The results in table 5 are taken from several test sequences. According to the results in this table, RDOQ is the most complicated algorithm accounting for the trellis search used, and the fixed-offset deadzone HDQ is the simplest algorithm due to coefficient-wise simple rounding processing. In addition, we also calculate the complexity increment of Huang's method and the proposed algorithm for each class sequences. T_D indicates the coding time of one frame by the Conventional Deadzone. T_X indicates the coding time of one frame by the Huang's method (T_H) or the proposed method (T_P). ΔT_X indicates the complexity increment of Huang's method or the proposed compared with Deadzone, which is expressed as follows.

$$\Delta T_X = \frac{T_X - T_D}{T_D} \times 100, X = HorP \tag{22}$$

As shown in Table 5,the proposed algorithm is slightly complex than the fixed-offset deadzone HDQ, with approximated $15\% - 33\%$ complexity increment. The increment is mainly contributed by initial online training, DCT distribution parameter estimation and update, and deadzone offset estimation. In addition, the proposed algorithm is slightly computation-efficient than the deadzone HDQ algorithm in [17] due to the offline trained deadzone offset model.

It should be mentioned that coefficient-wise processing is also adopted in the proposed algorithm. The flowchart of our algorithm in shown in Fig.13. As shown in this figure, DCT coefficients are quantized by HDQ at first. The quantization integer (ζ) and remainder(ξ) of HDQ rounding are obtained. The optimal offset is calculated according to the equation (21) with three parameters. And finally, the quantization results are ontained by adding the integer and $floor(\xi + \delta_{opt})$. Therefore, compared with the fixed-offset HDQ algorithm, the additional computation comes from two part, parameter Λ estimation and δ_{opt} calculation. Parameter Λ is estimated one time for one frame, online update only consume simple calculation. In addition, regardless of the initial online training, δ_{opt} are built offline and its calculation only needs simple function call as shown in equation (21) or tabulation shown in Fig.9 and 10, thereby it is moderate and trivial. As a result, the proposed algorithm is well-suited for parallel and pipelined hardware implementation.

The hardware friendliness of the proposed algorithm is meaning compared with the dynamic programming trellis-based SDQ and RDOQ. HDQ is well-suited for hardware implementation due to that coefficients in one block can be quantized simultaneously. Comparatively, trellis-based SDQ and RDOQ suffer from deadly inter-coefficient dependency, and coefficients in one block can only be processed in serial as analyzed in section I. This work proposes adaptive deadzone offset models offline, and the offsets are buffered in on-chip memory. With aid of the tabularized deadzone offsets, the proposed algorithm maintains the advantage of HDQ, i.e. hardware friendliness and coefficient-wise parallelism.

V. CONCLUSION

Sequential processing hinders soft-decision quantization (SDQ and RDOQ) from effective hardware implementations, whilst hard-decision quantization (HDQ) suffers from obvious rate distortion performance loss compared with SDQ. Based on statistical analysis and maximum a posterior estimation based modeling method, this paper proposes a content-adaptive deadzone HDQ algorithm to minimize the coding performance loss compared with RDOQ. An adaptive deadzone offset model is developed according to the quantization parameter, quantization remainder and coefficient-wise DCT distribution parameter. Simulation results verify that the effectiveness of the proposed adaptive HDQ algorithm. In comparison with conventional deadzone HDQ, it achieves 0.54% and 0.52% bit rate saving in two structure cases with nearly negligible complexity increment. In test dataset this work, in comparison with the RDOQ, achieves less than 0.55% and 0.54% bit rate increment on average in two structure cases. The proposed work is well-suited for hardwired video coding with the advantage of high-level

paralleling in HDQ thanks to the proposed coefficient-wise deadzone offset model.

ACKNOWLEDGMENT

The authors thank anonymous reviewers and editor for meaningful suggestions.

REFERENCES

[1] 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.

[2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[3] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[4] F. Bossen, B. Bross, K. Sühring, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, Dec. 2012.

[5] G. J. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1365–1374, Sep. 1996.

[6] V. Sheinin, A. Jagmohan, and D.-K. He, "On the operational rate-distortion performance of uniform scalar quantization-based Wyner–Ziv coding of Laplace–Markov sources," *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1225–1236, Nov. 2008.

[7] G. J. Sullivan and S. Sun, "On dead-zone plus uniform threshold scalar quantization," *Proc. SPIE*, vol. 5960, pp. 1041–1052, Jul. 2006.

[8] E.-H. Yang and X. Yu, "Rate distortion optimization for H.264 interframe coding: A general framework and algorithms," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1774–1784, Jul. 2007.

[9] G. J. Sullivan, *Adaptive Quantization Encoding Technique Using an Equal Expected-Value Rule*, document JVT-N011, ISO/IEC JTC1/SC29/WG11, ITU-T Q.6/SG16, 2005.

[10] E.-H. Yang and X. Yu, "Soft decision quantization for H.264 with main profile compatibility," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 1, pp. 122–127, Jan. 2009.

[11] E.-H. Yang, Z. Zhang, and T. Berger, "Fixed-slope universal lossy data compression," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1465–1476, Sep. 1997.

[12] J. Wen, M. Luttrell, and J. Villasenor, "Trellis-based R-D optimal quantization in H.263+," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1431–1434, Aug. 2000.

[13] M. Karczewicz, Y. Ye, and I. Chong, *Rate Distortion Optimized Quantization*, document VCEG-AH21, ITU-T SG16/Q.6, 12-C13, Antalya, Turkey, Jan. 2008.

[14] H. B. Yin, E. H. Yang, X. Yu, and Z. Xia, "Fast soft decision quantization with adaptive preselection and dynamic trellis graph," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1362–1375, Aug. 2015.

[15] L. Liu and A. Yourapis, *Rate Distortion Optimized Quantization in the JM Reference Software*, document JVT-AA027, Apr.2008.

[16] J. Cui, S. Wang, S. Wang, X. Zhang, S. Ma, and W. Gao, "Hybrid laplace distribution-based low complexity rate-distortion optimized quantization," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3802–3816, Aug. 2017.

[17] T.-Y. Huang and H. H. Chen, "Efficient quantization based on rate–distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 6, pp. 1099–1106, Jun. 2016.

[18] H. Lee, S. Yang, Y. Park, and B. Jeon, "Fast quantization method with simplified rate–distortion optimized quantization for an HEVC encoder," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 107–116, Jun. 2016.

[19] K. Lee, H.-J. Lee, J. Kim, and Y. Choi, "A novel algorithm for zero block detection in high efficiency video coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1124–1134, Dec. 2013.

[20] H. Wang and S. Kwong, "Hybrid model to detect zero quantized DCT coefficients in H.264," *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 728–735, Jun. 2007.

[21] B. Lee, J. Jung, and M. Kim, "An all-zero block detection scheme for low-complexity HEVC encoders," *IEEE Trans. Multimedia*, vol. 18, no. 7, pp. 1257–1268, Jul. 2016.

[22] H. Fan, R. Wang, L. Ding, X. Xie, H. Jia, and W. Gao, "Hybrid zero block detection for high efficiency video coding," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 537–543, Mar. 2016.

[23] J. T. Wen, M. Xiao, J. Chen, P. Tao, and C. Wang, "Fast rate distortion optimized quantization for H.264/AVC," in *Proc. Data Compress. Conf.*, Mar. 2010, p. 557.

[24] T.-Y. Huang, C. K. Kao, and H. H. Chen, "Acceleration of rate-distortion optimized quantization for H.264/AVC," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2013, pp. 473–476.

[25] J. He and F. Yang, "High-speed implementation of rate-distortion optimized quantization for H.264/AVC," *J. Signal Image Video Process.*, vol. 9, no. 3, pp. 543–551, 2015.

[26] J. Sole, R. Joshi, T. Ji, M. Karczewicz, G. Clare, and A. Duenas, "Transform coefficient coding in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1765–1777, Dec. 2012.

[27] V. Sze and M. Budagavi, "High throughput CABAC entropy coding in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1778–1791, Dec. 2012.

[28] K. Won, J. Yang, and B. Jeon, "Fast CABAC rate estimation for H.264/AVC mode decision," *Electron. Lett.*, vol. 48, no. 19, pp. 1201–1203, Sep. 2012.

[29] J. Hahm and C.-M. Kyung, "Efficient CABAC rate estimation for H.264/AVC mode decision," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 310–316, Feb. 2010.

[30] S. Choi and S.-I. Chae, "Comparison of CABAC rate estimation models for HEVC rate distortion optimisation," *Electron. Lett.*, vol. 50, no. 6, pp. 441–442, Mar. 2014.

[31] F. Bossen, *CE1: Table-Based Bit Estimation for CABAC*, document ITU-T Q.6/SG16, JCTVCG763, ITU-T, Geneva, Switzerland, 2011.

[32] H. Yin, H. Jia, J. Zhou, and Z. Gao, "Survey on algorithm and VLSI architecture for MPEG-like video coder," *J. Signal Process. Syst.*, vol. 88, no. 3, pp. 357–410, 2017.

[33] H. Wang, H. Yin, and Y. Shen, "A novel hard-decision quantization algorithm based on adaptive deadzone offset model," in *Proc. Pacific Rim Conf. Multimedia*, vol. 9917, 2016, pp. 335–345.

[34] (2003). *JVT Reference Software*. [Online]. Available: http://iphome.hhi.de/suehring/tml/

[35] (2013). *High Efficiency Video Coding Test Model Software HM 16.1.1*. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware

[36] J. Sun, Y. Duan, J. Li, J. Liu, and Z. Guo, "Rate-distortion analysis of dead-zone plus uniform threshold scalar quantization and its application—Part I: Fundamental theory," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 202–214, Jan. 2013.

[37] H. Yin, H. Wang, X. Wang, and Z. Xia, "Coefficient-wise deadzone hard-decision quantizer with adaptive rounding offset model," in *Proc. Data Compress. Conf.*, Mar./Apr. 2016, p. 638.

[38] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *Proc. Int. Conf. Image Process.*, vol. 3, Oct. 2001, pp. 542–545.

[39] M. Xiao, J. Wen, J. Chen, and P. Tao, "AN efficient algorithm for joint QP and quantization optimization for H.264/AVC," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 1233–1236.

[40] X. Li, N. Oertel, A. Hutter, and A. Kaup, "Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 193–205, Feb. 2009.

[41] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-L1100, ITU-T/ISO/IEC JCT-VC, Geneve, Switzerland, 2013.

[42] G. Bjontegaard, *Calculation of Average PSNR Differences Between RD-Curves*, document VCEG-M33, ISO/IEC MPEG ITU-T VCEG, JVT, Mar. 2001.

[43] G. Sullivan and K. Minoo, *Objective Quality Metric and Alternative Methods for Measuring Coding Efficiency*, document JCTVC-H0012, ITU-T/ISO/IEC, JCT-VC, San Jose, CA, USA, Feb. 2012, pp. 1–10.