


A Novel Hard-Decision Quantization Algorithm Based on Adaptive Deadzone Offset Model

Hongkui Wang, Haibing Yin , and Ye Shen

China Jiliang University, No. 258, Xueyuan Street, Hangzhou 310018, People's Republic of China
1550661343@qq.com, haibingyin@163.com, 04A0303080@cjlu.edu.cn

Abstract. In video encoder, the soft-decision quantization (SDQ) achieves superior coding performance however suffering from deadly sequential processing dependency. Comparatively, deadzone hard-decision quantization (HDQ) is dependency-free and suitable for hardwired parallel processing, however suffering from non-negligible coding performance degradation. In this paper, a content-adaptive HDQ algorithm is proposed by employing an adaptive deadzone offset. The contributions of this paper are characterized as follows. On one hand, this work applies offline statistic analysis, Bayes method, to explore the distribution characteristics of the desired deadzone offsets obtained from huge amounts of samples by fully simulating the behaviour of SDQ, and then derives adaptive deadzone offset model by maximizing the probability of right judgment of offset-induced rounding in HDQ. On the other hand, the deadzone offset model is constructed heuristically as functions of quantization step size, the distribution parameter of DCT coefficients and the number of possible significant coefficients prior to the current coefficient in the block. Simulation results verify that the proposed adaptive HDQ algorithm, in comparison with fixed-offset HDQ, achieves 0.08836 dB PSNR increment and 3.097 % bit rate saving with almost negligible complexity increase. In addition, this work, in comparison with the SDQ, achieves less than 0.03921 dB PSNR loss and 1.51 % bit rate increment. The proposed HDQ is well-suited for hardware encoder design.

Keywords: Video coding · Soft Decision Quantization · Hard Decision Quantization

1 Introduction

In MPEG-like hybrid video encoders, quantization plays an overwhelmingly important role in coding rate-distortion (RD) performance. It not only determines the quantization distortion, but also has prominent impacts on the consumed coding rate. Since video coding standards only define the inverse quantization, many research works have explored how to efficiently quantize discrete cosine transform (DCT) coefficients while remaining compliant with respective video coding standards [1].

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 [1]. In USQ + DZ, fixed rounding offsets

are usually employed and determined in a heuristic and empirical way. Coefficient-independent rounding based quantization in USQ and USQ + DZ is the so-called hard-decision quantization (HDQ), in which correlation among adjacent coefficients and their effects on quantization are not considered. The coefficient-wise processing in HDQ makes it friendly to hardwired video coder with parallel implementation. However, HDQ even deadzone HDQ both suffer from non-negligible rate distortion performance loss.

Soft decision quantization (SDQ) is a better alternative achieving superior RD performance contributed by full utilization of inter-coefficient correlation. A popular SDQ implementation is Viterbi trellis search [3], which was implemented for H.263+ [2] and for H.264/AVC [3]. However, running dynamic programming over the full trellis graph is computationally expensive. To get around this, in H.264/AVC and HEVC reference softwares JM and HM, a simplified suboptimal SDQ, called rate distortion optimized quantization (RDOQ) [4], was adopted. RDOQ is a simplified version of SDQ, implemented by employing dynamic programming in a similar way. SDQ achieves superior coding RD performance, approximately 6 ~ 8% bit rate saving as opposed to HDQ. In SDQ and RDOQ, multiple candidate results of quantization are competed and chosen using rate distortion optimization. As a result, heavy computation burden is one major challenge for SDQ and RDOQ. Moreover, dynamic programming based path search in SDQ results in severe sequential dependency, and serial processing in CABAC aggravates the data dependency in CABAC based SDQ [5].

Accounting for this issue, some literatures had made meaningful explorations to alleviate the computation burden of the SDQ [5–7]. These works decrease the computation of SDQ by decreasing the candidates of quantized results [5], employing fast computation for rate distortion of candidate coefficient levels [6, 7], and using fast bit rate evaluation [7]. These methods alleviate the computation burden in SDQ to certain extent, however they still suffer from some sequential dependency, or mainly designed for soft-targeted video coder optimization. In comparison, data dependency does not appear in the HDQ, such as the prevalent deadzone based HDQ. Coefficient level parallel processing can be achieved in coefficient-wise HDQ with obviously increased throughput by employing hardwired pipelining. As a result, HDQ is well-suited for hardwired video coder in terms of satisfactory throughput efficiency. Unfortunately, there is a nontrivial rate distortion performance gap between HDQ and SDQ.

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 simulating the behavior pattern of SDQ. On one hand, the distribution characteristics of DCT coefficients have great influence on quantization results, and thus DCT distribution parameter and quantization parameter are taken as consideration factor from the viewpoint of macro level. On the other hand, context modelling is used in CABAC and the number of possible significant coefficients plays important role in determining the quantization results. From the viewpoint of micro level manipulation, the deadzone offset is supposed to be tuned taking inter-coefficient influence into consideration according to the number of the possible significant coefficients in the block.

According to the above analysis, this paper aims to optimize rounding offset model for deadzone HDQ to improve the coding RD performance. Bayes classification method is used to derive the coefficient-wise deadzone offset model which is described as

functions of these three parameters, i.e. the quantization parameter, the parameter of component-wise DCT coefficients, and the number of possible significant coefficients prior to the current coefficient. In addition, the behaviour pattern of SDQ is analyzed and used as guidance for offset modelling to improve the RD performance of the proposed offset model based HDQ. This algorithm that proposed in this paper is well-suited for hardware coder design and achieves superior RD performance compared with deadzone HDQ thanks to considering the inter-coefficient correlation by offline methods to analyze the number of significant coefficients in the block.

The rest of this paper is organized as follows. Problem formulation is given in Sect. 2. The proposed HDQ algorithm is given in Sect. 3. Section 4 gives the experimental results. Finally, Sect. 5 concludes the paper.

2 Background and Problem Formulation

2.1 Difference Between Deadzone HDQ and SDQ

In CABAC-based SDQ, the output level of one coefficient not only depends on the levels of the anterior coefficients, i.e. backward dependency, but has influence on the following coefficients, i.e. forward influence. Intrinsically, dynamic programming such as Viterbi search is desired to track the inter-coefficient dependency in SDQ [5]. There are multiple sequentially scanned coefficients in one block, and each coefficient is described as a trellis stage in the graph [5]. There are multiple candidate quantized levels to be checked at each stage in SDQ, and they are described as candidate context states. The sequential coefficients and their candidate quantized levels form a trellis graph, and so the SDQ is actually a problem of searching for a path in the graph with minimum coding cost. The optimal path is composed by multiple adjacent branches, one survivor branch at one stage. The SDQ algorithm achieves superior coding performance, approximately 6 ~ 8 % bit rate saving, at the cost of high dependency caused by Viterbi algorithm and CABAC.

The HDQ algorithm is coefficient-wise based on memoryless source assumption, i.e. there is no dependency among adjacent coefficients. As a result, HDQ is well-suited for parallelism implementation. Compared with USQ, the USQ + DZ achieves considerable RD performance improvement by employing the statistical characteristics of entropy coding [1]. However, fixed offsets, 1/3 and 1/6 for intra and inter modes, are used in deadzone HDQ in H.264 JM and HEVC HM codecs [8]. Memoryless source assumption does not hold for the context based entropy coding, such as CABAC. It means that the fixed-offset HDQ is not fully optimized compared with optimal SDQ in terms of rate distortion optimization.

It is meaningful to excavate the inner characteristics of SDQ as guidance to propose adaptive offset model for new deadzone HDQ. The goal is to approach the RD performance of the SDQ and maintain the advantage of coefficient-wise processing in HDQ. The source distribution parameter and inter-coefficient influence will be taken into consideration.

2.2 Challenge in Adaptive Coefficient-wise HDQ

A deadzone offset δ is employed to adjust the quantization result ϕ in deadzone HDQ, and that can be described as follows.

$$\phi = \text{floor}\left(\frac{|u|}{q} + \delta\right) \quad (1)$$

Here, q is the quantization step size determined by the quantization parameter Q_p . In H.264/AVC and HEVC, q is equal to $\text{pow}(2, (Q_p-4)/6)$. floor is the rounding operator. u is the DCT coefficients to be quantized. $| \cdot |$ is the absolute value of the operation. There are several factors that should be considered in designing optimal coefficient-wise deadzone HDQ.

Firstly, the deadzone offset δ in HDQ is supposed to be determined in a coefficient-wise way. In general, Laplacian distribution is used to model the DCT coefficient, and the probability density function is described as follows.

$$f_i = \frac{1}{2\Lambda_i} e^{-\frac{|u_i|}{\Lambda_i}} \quad \text{and} \quad \Lambda_i = \frac{\sigma_i}{\sqrt{2}} = \frac{1}{n} \sum_{j=1}^n |u_{ij}| \quad (2)$$

Λ_i and σ_i are the model parameter and the standard deviation of the i th frequency component. u_{ij} is the DCT coefficient and u_{ij} is the j th DCT coefficient of the i th frequency component. Based on Laplacian DCT distribution model, the deadzone offset δ of HDQ is typically determined as follows using rate distortion optimization [1].

$$\delta = \frac{q}{2} - \frac{\lambda}{2 \ln 2 \times \Lambda} \quad (3)$$

Here, λ is Lagrange multiplier and equals to $\ln 2(\text{power}(q-\delta, 2) - \text{power}(\delta, 2))/q$ in [1]. However, there is an egg-and-chicken problem because the parameters λ and δ are dependent with each other. The coefficient-wise model in (3) is built in a macroscopic way based on statistical analysis. However, the SDQ algorithm manipulates the quantization result in a microscopic way, specifically according to the probabilities of the contexts in CABAC. Moreover, based on λ equals to $\ln 2(\text{power}(q-\delta, 2) - \text{power}(\delta, 2))/q$, coefficient level solutions are theoretically desired for Eq. (3), which is not easy to be solved.

Thirdly, the adaptive deadzone offset in Eq. (3), derived from coefficient-level models without considering inter-coefficient influence, will unavoidably suffer from RD performance degradation. In CABAC, coding bit of a quantization coefficient is determined by the probability state of the context, which is modelled according to the numbers of coefficients with intensity equal to 1 and larger than 1, i.e. Numeq1 and NumLg1 , prior to the current coefficient. Under the criterion of rate distortion optimization, SDQ considers the inter-coefficient correlation and adjusts the quantization coefficient level. The code rate of the quantization coefficient is related to the number of nonzero (significant) quantization coefficients which are in current block and in the adjacent block.

Therefore, this paper take this factor into consideration in a micro way, i.e. employing the number of significant coefficients in one block in determining the dead-zone offset and quantization results. However, the number of significant coefficient of a certain coefficient is not available before the SDQ algorithm finish the trellis search. Thus, it is not possible to measure the accurate number of significant coefficient in deadzone HDQ. By striking a compromise, we define the number of possible significant coefficients for the i -th coefficient η_i according to the HDQ quantization results as follows.

$$\eta_i = \sum_{j=i+1}^N \varpi(|u_j| - \frac{q}{2}) \text{ and } \varpi(x) = \begin{cases} 1; & x > 0 \\ 0; & x \leq 0 \end{cases} \quad (4)$$

When the amplitude of the DCT coefficient is larger than $q/2$, the coefficient is considered as a possible significant coefficient. η_i is the number of possible significant coefficients prior the current i -th coefficient. For all coefficients on one block, their η_i can be estimated in parallel due to the coefficient-wise processing in HDQ as shown in (4). Suppose N is the number of the last DCT coefficient in the adjacent block, so η_i is an integer ranging from 0 to $N-i$.

Here we first quantitatively evaluate the degree of which η_i affect the quantization result of the current i -th coefficient. We use $\Phi_{HDQ(u)}$ and $\Phi_{SDQ(u)}$ to distinguish the quantization amplitude of u in HDQ and SDQ respectively. There are two possible situations according to the condition whether $\Phi_{HDQ(u)}$ is equal with $\Phi_{SDQ(u)}$ or not, i.e. $\Phi_{HDQ} = \Phi_{SDQ}$ and $\Phi_{HDQ} \neq \Phi_{SDQ}$. Taking 4×4 transform block as an example, the statistical distribution of η_i of sequences in the Table 1 is counted respectively. Figure 1 shows these sequences' average results of the statistical distribution of η_i in the case of two cases. We can make comparison between the η_i results of two cases, i.e. $\Phi_{HDQ} = \Phi_{SDQ}$ and $\Phi_{HDQ} \neq \Phi_{SDQ}$, as shown and in Fig. 1. The result shows that the distribution of η_i is more dispersed when

Table 1. The BD-PSNR loss of the deadzone HDQ algorithms using fixed-offset and the proposed adaptive offset compared with optimal SDQ

Sequences	BD-PSNR (dB)		BD-RATE (%)		Sequences	BD-PSNR (dB)		BD-RATE (%)	
	fixed-offset	Proposed	fixed-offset	Proposed		fixed-offset	Proposed	fixed-offset	Proposed
City	-0.1123	-0.0809	2.7651	2.0926	SlideShow	-0.2011	-0.0951	3.1697	1.5151
Harbour	-0.1505	-0.1145	3.2987	2.4996	720p Average	-0.1146	-0.0429	2.9833	1.1345
Crew	-0.1396	-0.0442	1.3462	0.7649	Proposed vs fixed-offset	0.0717		-1.8488	
ICE	-0.0711	-0.0350	1.9779	0.9418	Kimono1	-0.1675	-0.0186	6.1667	0.8812
Soccer	-0.1675	-0.0807	3.7417	1.7785	ParkScene	-0.0647	-0.0468	2.0155	1.4357
D1 Average	-0.1282	-0.0710	2.6259	1.6155	riverbed	-0.1542	-0.0093	3.1231	0.2450
Proposed vs fixed-offset	0.0572		-1.2968		Basketball	-0.0881	-0.0542	4.1461	2.6046
Cyclists	-0.0627	-0.0353	2.3778	1.3271	Cactus	-0.0892	-0.0189	4.4226	0.3890
Harbour	-0.1013	-0.0308	2.4938	0.9619	sunflower	-0.2016	-0.0874	7.76895	3.5057
Optis	-0.0759	-0.0079	2.6692	0.3354	1080p Average	-0.1275	-0.0392	4.6072	1.5102
Raven	-0.132	-0.0455	4.2062	1.5328	Proposed vs fixed-offset	0.08836		-3.0969	

$\Phi_{HDQ} = \Phi_{SDQ}$, while η_i are mainly concentrated in the vicinity of zero comparatively in the case of $\Phi_{HDQ} \neq \Phi_{SDQ}$. This statistical differences of η_i in two opposite cases give us the insight that η_i can be employed to aid in deriving deadzone offset model in terms of simulating SDQ. This work will adjust the deadzone offset δ according to the actual distribution of η_i to simulate the SDQ decision mechanism as far as possible.

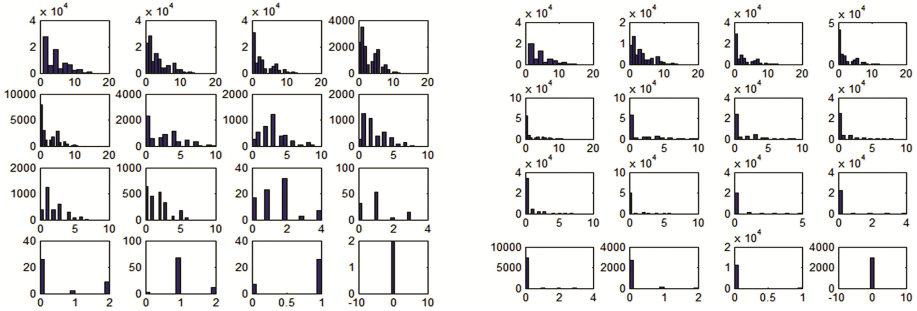


Fig. 1. The histogram results of η_i in the case of $\Phi_{HDQ} = \Phi_{SDQ}$ and $\Phi_{HDQ} \neq \Phi_{SDQ}$

3 Improved HDQ with the Proposed Adaptive Deadzone Offset

3.1 Heuristic Deadzone Offset Modelling

As analyzed above, deadzone offset modelling is supposed to be built adaptively for deadzone HDQ. Instead of RD model based derivation method shown in Eq. (3), this work attempts to estimate optimal deadzone offset model by simulating the behavior of SDQ and construct an adaptive deadzone offset model based on statistical analysis method.

As shown in Fig. 2, statistical analysis and heuristic method for model derivation is employed. We use the classification method in II-C to distinguish two kinds of situations. The “inlier” and “outlier” which represent DCT coefficients samples of two categories respectively, are collected for offline deadzone offset modeling. The “inlier” samples are the DCT coefficients in the case of $\Phi_{HDQ} = \Phi_{SDQ}$; the “outlier” samples are the DCT coefficients in the case of $\Phi_{HDQ} \neq \Phi_{SDQ}$. Taking the inter-coefficient correlation into consideration as analyzed in (3), the deadzone offset δ is related with the quantization parameter Q_p , the Laplacian distribution parameter Λ and the number of possible significant coefficients prior to the current coefficient, i.e. η_i . As a result, the parameter (Q_p, Λ, η) combination samples of the “inlier” and “outlier” coefficients are collected for statistical analysis. In addition, the “inlier” offset range ($\delta_{min1}, \delta_{max1}$) and the “outlier” offset range ($\delta_{min2}, \delta_{max2}$), in which the resulting results of HDQ are equal to those of SDQ of two kinds of samples, are also recorded simultaneously. These statistic samples will be used for off-line offset modeling.

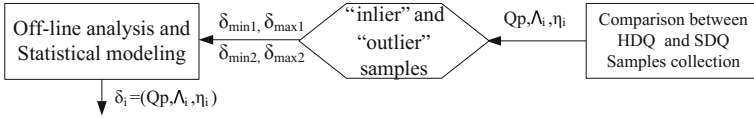


Fig. 2. Heuristic deadzone offset model derivation.

3.2 Analysis on Optimal Offset Distribution

When the HDQ fails to track the optimality of SDQ, the quantization results of HDQ will be different from those of SDQ, i.e. $\Phi_{HDQ(u)} \neq \Phi_{SDQ(u)}$. In these samples, we find that the probability of $\Phi_{HDQ(u)} = \Phi_{SDQ(u)} + 1$ is very close to 1, and this phenomenon can be explained that smaller quantization level in SDQ, with changed level intensity equal to 1, will result in predominate rate saving, which is larger than the increased distortion. The algorithm design of adaptive deadzone HDQ is just to find adaptive deadzone offset (δ_{best}) which is suitable for both of two kinds of samples.

The method of deriving the ranges of the possibly reasonable offsets is analyzed as follows. A well-designed offset δ for deadzone HDQ may achieve right classification with result identical with SDQ, or wrong classification with result differing from SDQ. δ_{best} will be determined from a great amount of samples using statistical analysis method, i.e. Bayes method. Intuitively, we estimate appropriate deadzone offset range that can make HDQ achieving the identical result as SDQ. The suitable deadzone offset ranges of two cases are estimated respectively, and their upper and lower bounds of offset ranges of two kinds of samples, i.e. $(\delta_{min1}, \delta_{max1})$ and $(\delta_{min2}, \delta_{max2})$, are estimated as follows.

$$\begin{aligned}
 (\delta_{min1}, \delta_{max1}) &= \arg(\underset{\delta_{best}}{\text{floor}}(\frac{|u|}{q} + \delta_{best}) = \Phi_{HDQ(u)}) \\
 (\delta_{min2}, \delta_{max2}) &= \arg(\underset{\delta_{best}}{\text{floor}}(\frac{|u|}{q} + \delta_{best}) = \Phi_{HDQ(u)} - 1)
 \end{aligned}
 \tag{5}$$

Here, $\Phi_{HDQ(u)} = \text{floor}(|u|/q + 0.5)$ is HDQ quantization intensity. As for the “inlier” samples, we get the possible range of the optimal δ under the constraint that $\text{floor}(|u|/q + \delta_{best})$ is equal to $\Phi_{HDQ(u)}$. As for the “outlier” samples, we get the possible range of the optimal δ under the constraint that $\text{floor}(|u|/q + \delta_{best})$ is equal to $\Phi_{HDQ(u)} - 1$. In a word, we estimate the possible ranges of the optimal δ under the constraint of $\Phi_{HDQ(u)} = \Phi_{SDQ(u)}$ for both kinds of samples.

Offline statistic analysis is carried out using a great amount of two kinds of samples and their offset ranges. Histogram based non-parametric analysis is employed to estimate these two probability density function curves. The maximal possible solution range for $(\delta_{min}, \delta_{max})$, i.e. $(0, 1)$, is partitioned into N segments, and the actual range $(\delta_{min}, \delta_{max})$ of all samples are compared for grouping and classification respectively. Then, ranges $(\delta_{min}, \delta_{max})$ of all samples are compared for classification respectively. If one segment is within range $(\delta_{min}, \delta_{max})$, its histogram count $\theta_y(k)$ is increased by $\zeta(k)$. $\theta_y(k)$ is expressed as follows.

$$\begin{aligned} \theta_y(k) &= \theta_y(k) + \zeta(k); & \text{if } \left(\frac{k-1}{N}, \frac{k}{N}\right) \in (\delta_{\min_y}, \delta_{\max_y}) \\ \theta_y(k) &= \theta_y(k); & \text{otherwise} \end{aligned} \quad y = 1 \text{ or } 2 \tag{6}$$

In this work, different weight is used for histogram estimation accounting for statistical characteristics. That is, different weight $\zeta(k)$ is used for different each subsection for δ_{best} histogram statistical analysis. Gaussian function is used for modeling the weight of each sub section ($\zeta(k)$). Gaussian function based $\zeta(k)$ is employed and expressed as follows.

$$\zeta(k) = \frac{1}{\sigma_1 \times \sqrt{2 \times \pi}} \times e^{\left(\frac{k - \mu_1}{2 \times \sigma_1^2}\right)^2} \quad \text{and } \mu_1 = \frac{\delta_{\max} + \delta_{\min}}{2}, \sigma_1 = \frac{\delta_{\max} - \delta_{\min}}{\alpha} \tag{7}$$

Here, α is equal to 6. The segment-wise histogram results of δ_{best} , $\text{cnt}_y(\text{Qp}, \Lambda)$ are obtained independently in the case of different Qp and Λ . $y = 1$ and $y = 2$ correspond to the ‘‘inlier’’ and ‘‘outlier’’ samples. It is well-known that Λ is related with the coefficient position index (i). One example of $\text{cnt}_y(\text{Qp}, \Lambda(i))$ of different coefficients with two kinds of samples is shown in Fig. 3.

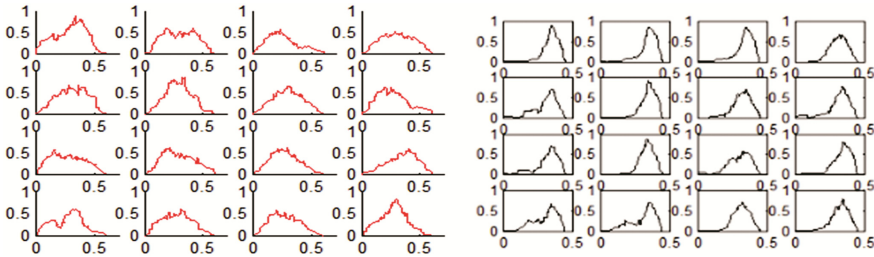


Fig. 3. The histogram results of possible δ_{best} $\text{cnt}_1(\text{Qp}, \Lambda)$ and $\text{cnt}_2(\text{Qp}, \Lambda)$

3.3 The Adaptive Deadzone Offset Model

According to the statistical samples, we build an adaptive model $\delta_i = (\text{Qp}, \Lambda_i)$ using heuristic method for deadzone offset. The model is constructed by maximizing the positive judgment probability and minimizing the probability of wrong judgment. In fact, $\theta_1(k)$ and $\theta_2(k)$ actually reflect the right classification probabilities of the ‘‘inlier’’ and ‘‘outlier’’ coefficients. We perform normalization for the histogram results, $\theta_y(k)$, to derive the condition probability p_y , i.e. $p_1(\delta)$ and $p_2(\delta)$ respectively. Therefore, δ_{best} can be determined by taking the peak value of the weighted histogram as shown in Eq. (8). The schematic diagram and the actual statistics results are shown respectively in Fig. 4.

$$\delta_i = \arg \max_{\delta_i(\Lambda_i, \text{Qp})} \{p_1(\delta_i) + p_2(\delta_i)\} \tag{8}$$

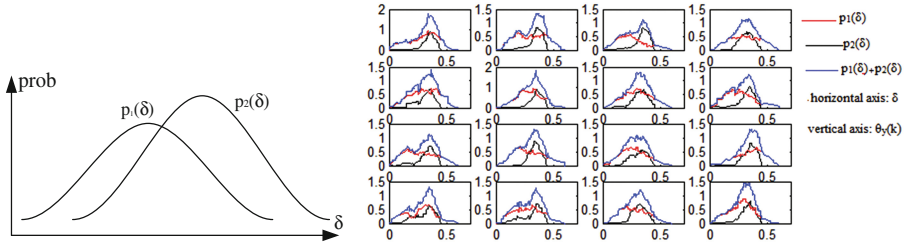


Fig. 4. The sketch map of δ_{best} and the statistical results of δ_{best}

On the basis of the above model, we built a factor $\phi(\eta_i)$ which can adjust the deadzone offset according to the number of possible significant coefficients in the adjacent block. Here, i is the coefficient position index and the factor $\phi(\eta_i)$ is expressed as follows.

$$\phi(\eta_i) = \beta \times \arctan(\eta_i - \frac{\eta_{imax}}{\gamma}) \quad \text{and} \quad \eta_{imax} = N - i \tag{9}$$

Here, N is the number of the last DCT coefficient in the adjacent block. η_{imax} is equal to $N-i$. We had evaluated the RD performance in the cases of different combinations between (β, γ) . The simulation results indicate that superior RD performance appear when β is equal to -0.03 and γ is equal to 3 , as shown in Fig. 5. Therefore, the adaptive deadzone offset model with η_i is expressed as follows.

$$\delta_i = \arg \max_{\delta_i(\Lambda, Qp)} \{p_1(\delta_i) + p_2(\delta_i)\} + \phi(\eta_i) \tag{10}$$

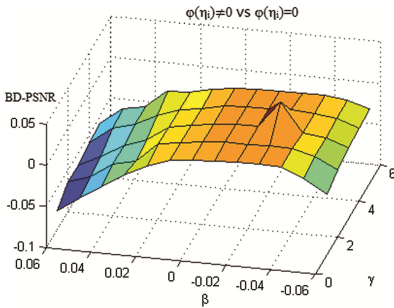


Fig. 5. The RD performance of different combinations between β and γ

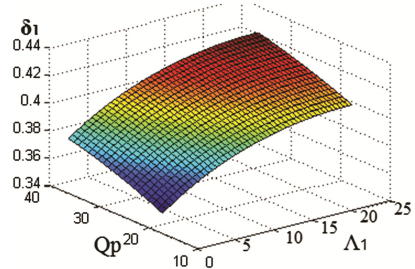


Fig. 6. The surf results of δ_{best}

According to the method of modeling above, in the case of $\eta_1 = 10$, one example of the surf results of δ_1 , in the case of different combinations between Qp and Λ_1 , is shown in Fig. 6.

4 Experimental Results

The proposed adaptive offset model for deadzone HDQ is verified in H.264 and H.265 standards. The fixed-offset HDQ algorithm and the optimal SDQ are taken as the performance comparison anchor. These quantization algorithms are applied in both the final mode coding and the rate distortion optimized mode decision loop. Standard D1, 720p, and 1080p format video sequences are used for simulation. Rate control is turned off, and the quantization parameters 22, 27, 32 and 37 are used for simulation, covering low, medium and high bit rate applications. IPBBPBB GOP structure is used, and 100 frames are tested for all resolution video sequences. The PSNR degradation (BD-PSNR) and rate increment percentage (BD-RATE) are used for performance comparison [8].

The rate distortion curves of the 1080p BasketballDrive sequence are taken as example shown in Fig. 7. The anchor optimal SDQ, the fixed-offset deadzone HDQ, and the proposed algorithm are compared. In addition, Table 1 gives the detailed BD-PSNR and BD-RATE results [8]. Relatively, larger RD performance improvement is observed in higher resolution video sequences. Intensive results show that the proposed algorithm only has 0.03921 dB BD-PSNR loss on average, with 1.51 % average BD-RATE increment, in comparison with the SDQ algorithm in the case of 1080p sequences. In addition, the proposed algorithm achieves 0.08836 dB BD-PSNR improvement on average, with 3.097 % average rate saving (BD-RATE), in comparison with the fixed-offset deadzone HDQ algorithm. The proposed adaptive offset HDQ algorithm is considerably superior than the fixed-offset HDQ, and has close performance with the optimal SDQ algorithm. Compared to the SDQ algorithm, the proposed algorithm has much smaller complexity and is well-suited for hardwired video coder in terms of satisfactory throughput efficiency.

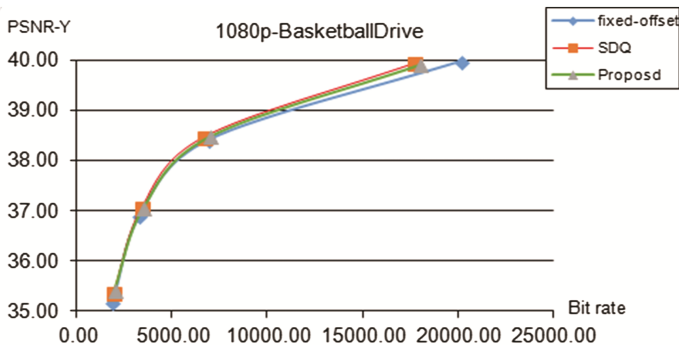


Fig. 7. The RD performance of three kinds of quantization algorithms.

As for complexity, the additional computation of the proposed algorithm, in comparison with the fixed-offset HDQ algorithm, is just simple function call as shown in Eq. (10) or tabulation shown in Fig. 6, so it is almost ignorable.

5 Conclusions

Sequential processing hinders soft-decision quantization (SDQ) from effective hardware implementations, while hard-decision quantization (HDQ) suffers from obvious coding performance loss compared with SDQ. Based on statistics analysis and heuristic modeling, this paper proposes a content-adaptive deadzone quantizer to minimize the rate distortion performance difference between the deadzone HDQ and SDQ. An adaptive deadzone offset model is built according to the quantization parameter, the coefficient-wise DCT distribution parameter, and the number of possible significant coefficients in the block. Simulation results verify that the proposed adaptive HDQ algorithm, in comparison with fixed-offset HDQ, achieves 0.08836 dB PSNR increment and 3.097 % bit rate saving in 1080p sequences with almost negligible complexity increase. In addition, this work, in comparison with the SDQ, achieves less than 0.03921 dB PSNR loss and 1.51 % bit rate increment.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China under Grants 61572449, in part by the Zhejiang Natural Science Foundation under Grants LY15F020022 LY12F01011, and LY13H180011.

References

1. Sullivan, G.J.: Efficient scalar quantization of exponential and laplacian random variables. *IEEE Trans. Inf. Theor.* **42**, 1365–1374 (1996)
2. Wen, J., Luttrell, M., Villasenor, J.: Trellis-based R-D optimal quantization in H.263+. *IEEE Trans. Image Process.* **9**, 1431–1434 (2000)
3. Yang, E.-H., Yu, X.: Rate distortion optimization of H.264 with main profile compatibility. In: *Proceedings of IEEE International Symposium on Information Theory, Seattle, WA*, pp. 282–286, 9–14 July 2006
4. Karczewicz, M., Ye, Y., Chong, I.: Rate Distortion Optimized Quantization. ITU-T SG16/Q. 6 Doc. VCEG-AH21, Antalya, Turkey, 12–13 January 2008
5. Yin, H., Yang, E., Xiang, Y.: Fast soft decision quantization with adaptive preselection and dynamic trellis graph. *IEEE Trans. Circ. Syst. Video Technol.* **25**(8), 1362–1375 (2015)
6. Huang, T.-Y., Chen, H.H.: Efficient quantization based on rate-distortion optimization for video coding. *IEEE Trans. Circ. Syst. Video Technol.* **26**, 1099–1106 (2015). doi:[10.1109/TCSVT.2015.2444732](https://doi.org/10.1109/TCSVT.2015.2444732)
7. Lee, H., Yang, S., Park, Y., Jeon, B.: Fast quantization method with simplified rate-distortion optimized quantization for HEVC encoder. *IEEE Trans. Circ. Syst. Video Technol.* (2015). doi:[10.1109/TCSVT.2015.2450151](https://doi.org/10.1109/TCSVT.2015.2450151)
8. Sun, J., Duan, Y., Li, J., Liu, J., Guo, Z.: Rate-distortion analysis of dead-zone plus uniform threshold scalar quantization and its application—part I: fundamental theory. *IEEE Trans. Image Process.* **22**(1), 202–214 (2013)