Journal of Computers Vol. 29 No. 3, 2018, pp. 73-82
doi:10.3966/199115992018062903007

A Memory Source Model Researched on Video Coding Quantization Algorithm

Xin-Xiu Wei[1], Zhe-Lei Xia[1*], Hai-Bing Yin[1, 2], Hong-Kui Wang[1]

[1] School of Information Engineering, China Jiliang University, Hangzhou 310018, China
wxxlittleapple@qq.com, xia663618@163.com, 1550661343@qq.com

[2] School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
yinhb@cjlu.edu.cn

Received 17 November 2016; Revised 6 March 2017; Accepted 30 March 2017

Abstract. In video encoder, calculating speed of hard decision quantization (HDQ) is fast, however, the rate distortion performance is lost. Comparatively, the coding performance of best soft decision quantization (SDQ) is better but the calculation is complex. In order to solve the limitation of these two algorithms, this paper proposes a quantization algorithm of memory source model based on HDQ. The contributions of this paper as follows. Firstly, we puts forward to a framework of hard decision quantization algorithm based on adaptive threshold which takes rate consumption into account, heuristic threshold modeling. Secondly, give the method which can determine the model of rate threshold offline based on the principle of maximizing the probability of right judgment and minimizing the probability of wrong judgment. Finally, according to the coefficient of quantization position and the distribution of quantization parameter, an adaptive threshold model is constructed based on rate distortion optimization, and the optimal parameters of the model are determined. Experiments show that compared with fixed offset HDQ algorithm, HDQ with adaptive threshold based on source memory in this paper obtains significant performance improvements, BD-PSNR has 0.0964dB boost, BD-Rate has about 3.5723% bit rate savings. Compared to SDQ, the additional computational complexity of this paper is low in real time computation. Adaptive HDQ based on this model is very suitable for designing of hardware encoder architecture.

Keywords: context adaptation, HDQ, rate distortion optimization, SDQ, video coding

1 Introduction

Quantization plays a very important role in mixed-frame video encoder. It directly determines the quantization distortion and directly affects the code rate [1-2]. In early video codec, quantization of transform coefficients was performed using the uniform scalar quantization, USQ [1]. Later based on dead-zone linear quantization because of its high performance is widely used [2-3]. All of above collectively referred to HDQ [3]. These algorithms don't consider correlation of adjacent coefficients in the quantization block, and regarded as coefficients is independent of each other. The rate-distortion performance is not optimal, which is still room for improvement [3].

Some academics put forward to the SDQ [4].This algorithm consider the data dependency between coefficients, meanwhile, taking rate expending and coding distortion into account. Yang and Wang applied SDQ algorithm into researches, and proposed a optimization of SDQ algorithm which supports context adaptive variable length coding (CAVLC) and context-based adaptive binary arithmetic coding (CABAC) that two methods of entropy coding [5], obtained higher compression coding efficiency. However, dependence of coefficient makes algorithm has high computational complexity.

A simplified algorithm of SDQ algorithm, called Rate distortion optimization quantization, RDOQ [6-7]. In this algorithm, the computational complexity is still high, which is not easy to implement for

* Corresponding Author

hardware. Thus, while maintaining the low complexity of the algorithm and improving the rate distortion performance, researchers need to think and research [8-9].

Based on the research of [10] algorithm, this paper proposed a new model, which draw deeply out the inherent characteristics of SDQ algorithm through simulating the characteristics of optimal SDQ, introducing correlation between coefficients. Built on HDQ, this new model in paper puts forward to a HDQ algorithm which supports for CABAC and threshold adaptive. The model has a very important significance, which breaks the data dependency of SDQ algorithm. It can save more bit rate and realize finely quantifying at coefficient level besides enhancing rate distortion performance of HDQ. And overcome serialization of SDQ data processing, also adapt to hardware platform parallel processing.

2 Problem Formulation

2.1 Deadzone HDQ and RDO Based SDQ

The goal of video compression is to reduce the compression rate of the video stream on the premise of ensuring the quality of the video, however, relationship between the rate of coding output and the distortion after compression is mutually restricted and contradictory. Therefore, the research of quantization algorithm in video coding needs to be combined with the rate distortion optimization theory, and also balances the distortion and bit rate. So, optimization of video quantization algorithm can be used with the cost function J. Solving method using Lagrangian cost function as follows.

$$J = D + \lambda \times R \tag{1}$$

Hard-decision Quantization (HDQ). Deadzone offset δ is employed to adjust the quantization result z in deadzone HDQ, and that can be described as follows.

$$z = \text{floor}\left(\frac{|u|}{q} + \delta\right) \tag{2}$$

Here, δ is 1/2.

The quantization offset δ fixed shows the coefficients of HDQ algorithm is quantified independently without considering the dependence of coefficients. However in the encoder with CABAC entropy coding, the assumption of no memory signal source is not set up [5]. According to the rate distortion formula, this kind of algorithm is aimed at reducing the distortion of D, making J decreasing, to a great extent, it ignores the effect of rate R on quantization results and the cost function J.

Soft-decision Quantization (SDQ). SDQ algorithm first appeared in literature [12], later, Yang, etc. present algorithm optimization based on CABAC, SDQ, which applied rate distortion optimization and dynamic programming optimization to coefficient combined optimization in quantization block. This algorithm adopts Viterbi dynamic programming, whose each branch (path) in the network represents a specific coefficient quantization result. Rate distortion optimization quantization using formula (1) to calculate the cost of each path, select one of possible paths whose rate distortion (RD) cost minimization for the optimal path. Ye, etc. reduced the number of SDQ Viterbi grid paths, and proposed a simplified version of SDQ at the expense of small performance loss, RDOQ [4].

RDOQ [13] is embedded the bottom module of the RDO loop. Viterbi dynamic programming and context entropy encoding leads to serious serial dependence of hardware and too high computational complexity. This paper tried to study the rate model to estimate rate of coefficient [14] based on HDQ, which can reduce the optimal quantization complexity and achieve in hardware.

2.2 Our Previous Deadzone Offset Model

In our previous work [10], firstly we must determine the adaptive offset δ'_{opt} in order to adjust δ_k. Then simulate the behavior characteristics of SDQ algorithm, and estimate the optimal deadzone offset in literature [10], based on the dual constraints of maximum positive judgment and the minimum probability of false judgment.

When SDQ is not consistent with the results of HDQ, it adjusts HDQ offset, making adjusted HDQ results are in agreement with the SDQ quantization results and regarding this offset range as interval $1(\delta_{min1}, \delta_{max1})$, according to the following formula determined.

$$(\delta_{min1}, \delta_{max1}) = \arg_{\delta opt}\{0.85 \le \frac{\mathrm{mod}(u,q)}{q} + \delta_{opt} \le 1\} \quad (3)$$

When SDQ is consistent with the results of HDQ algorithm, it adjusts offset, which ensures that two quantitative results has its concordant offset range, regarded as the interval $2(\delta_{min2}, \delta_{max2})$.

$$(\delta_{min2}, \delta_{max2}) = \arg_{\delta opt}\{0 \le \frac{\mathrm{mod}(u,q)}{q} + \delta_{opt} \le 0.5\} \quad (4)$$

According to the intersection of interval1 and interval2 that offset is located on, it can be determined the optimum offset range. The offset is constructed as a function of the DCT [11] distribution parameter and the quantization parameter QP, which is used to quantify the DCT coefficients adaptively.

$$\delta(Qp,\Lambda) = \rho - \mu \times e^{(\beta \times \frac{(Qp-a)^2}{\xi_1} + \gamma \times \frac{(\Lambda-b)^2}{\xi_2})/\alpha} \quad (5)$$

Thereinto, QP >17.

I frame ρ=0.50+0.001*(Qp-17), μ=0.1, β=0.6, γ=0.6, α=128, $\xi_1=\xi_2$=32, α=35, b=50.

P/B frame ρ=0.48+0.001*(Qp-17), μ=0.1, β=0.6, γ=0.6, α=128, $\xi_1=\xi_2$=32, α=35, b=50.

In short, from the perspective of rate distortion, the model in literature [10] only considers that the quantization parameters QP, DCT parameters effect distortion D, which ignores rate of each quantized coefficients effects on the quantization results. The researching content of this paper based on literature [10], introduces the rate information which each position carries on in the block as the factor influencing the quantization, at the same time, adjusting to the offset in literature [10].

3 Offset Model Building and Parameter Selection

According to different quantization results, each position in quantization block carried it own message segment, which exist bit rate consumption in the encoding process. Consequently, the offset model based on deadzone HDQ, which added estimating the probable rate consumption when different results was quantified, is essential for choosing appropriate quantitative results.

RDO quantization algorithm is designed to minimize the RDO cost of the quantization results Z. It is well known that rounding HDQ algorithm (δ=0.5) gets the lowest quantization result, yet the final quantization result is z-1 in SDQ algorithm. This phenomenon can be attributed that SDQ whose quantization amplitude decreases compared to HDQ will lead to increased quantization distortion, at the same time to save the encoding bits $\Delta R_{saved}=R_z-R_{z-1}$. When the rate saving cost is greater than the distortion cost, z-1 is better from the perspective of rate distortion optimization.

Assuming that z and z-1 are the most close to two integer candidate results with HDQ quantization floating point results, it was observed that SDQ selection of these two results is almost as high as 99%. SDQ algorithm essentially can be understood as the two value of the decision problem, whose task is to select the RDO cost minimum result from z and z-1.Bit difference ΔR that encoding quantized amplitude z and z-1expends has a vital role in SDQ two value judgment results. This paper intends to study the quantitative effect of ΔR for SDQ final results, and use it as an important parameter to construct the dead zone migration model.

3.1 Heuristic Threshold Model

Because of SDQ using Vitby dynamic programming optimization, CABAC coding based on the assumption of active memory model, and the interaction between the coefficients of the block or CG coefficients, the context of the encoding amplitude of current coefficient is determined by the proximity factor. In this paper, we improve the HDQ algorithm to try to approximate the SDQ algorithm, although the result is not consistent with SDQ, leading to HDQ algorithm can not accurately estimate ΔR value of each coefficient. This paper intends to choose the context based on the results of HDQ, then estimate ΔR value of current coefficient, ΔR'.

In the paper, it is considered that the ΔR statistic distributions of the two types of samples, z and z-1, are different. Through the offline analysis of two types of samples ΔR, distribution histograms are shown

in Fig. 1 and Fig. 2 (probability density function). The Bayesian classification method is used to determine the ΔR threshold which can distinguish the quantization result of z and z-1 based on the minimum misjudgment constraint denoted as ΔR_{th}, as shown in Fig. 3. Obviously, the degree of deviation between the actually estimated rate-of-use savings ΔR' and the threshold value ΔRth, $\psi_{\Delta R} = \Delta R' - \Delta R_{th}$, which quantitatively measures the degree of coding bit savings. In this paper, relationship between $\psi_{\Delta R}$ and ideal dead-zone offset is analyzed statistically so as to construct the dead-zone offset.

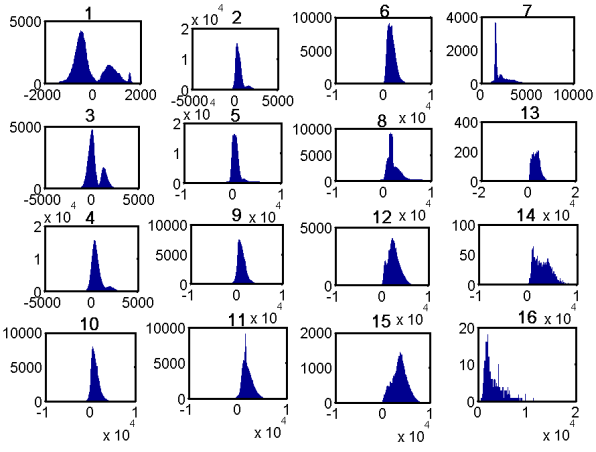

Fig. 1. SDQ is quantized to z-1

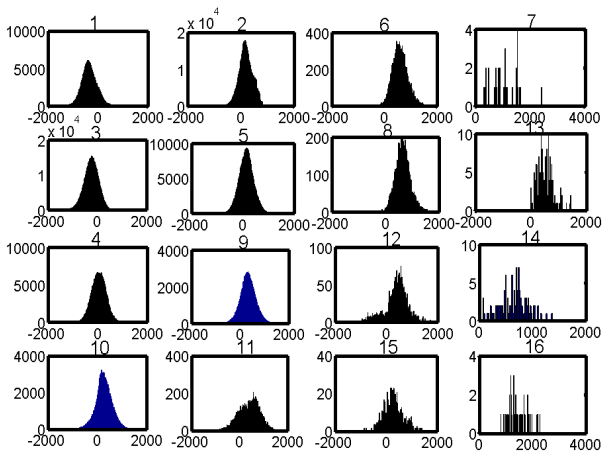

Fig. 2. SDQ is quantized to z

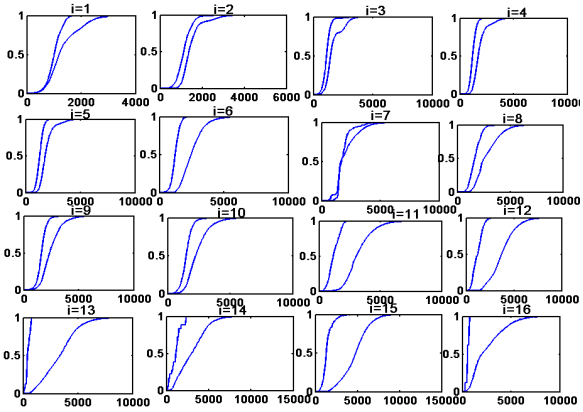

Fig. 3. CDF plot of the estimated threshold for each position ΔR4x4th in the 4x4 quantization block

$$\Delta R = R_z - R_{z-1} \tag{6}$$

$$\psi_{\Delta R} = \Delta R' - \Delta R_{th} \tag{7}$$

One of ΔR_{th} approximates SDQ quantization to z, which has positive contribution to SDQ quantizing to z, the other approximates SDQ quantization to z-1 and contributes positively to quantify level as z-1.

Experiments were performed on 1080p 6 sequences that 500 frames/per sequence. The remainder is controlled within 0.5 to 0.7, divided into four segments, [0.5, 0.55], [0.55, 0.6], [0.6, 0.65], [0.65, 0.7]. Qp range in [22, 40], the step size is Fig. 3. Finally, the ΔR_{th} corresponding to the 16 position coefficients of the 4x4 quantization block is measured as ΔR_{4x4th}. The scatter plot is plotted as a three-dimensional surface plot with x, y, and z coordinates as Qp, i, and ΔR4x4th, respectively, and the method of surface fitting is used to truncate singular points, the curve fitting curve of $\Delta R_{th} = f(Qp, i)$ as shown in Fig. 4.

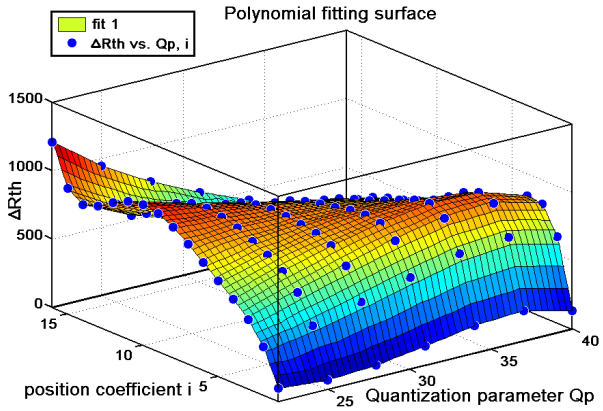

Fig. 4. ΔR_{th}= f (Qp, i) polynomial fitting surface

3.2 ΔR_{th} Threshold Analysis

From Fig. 3 and Fig. 5, the sampling point data quantity decreases with position of coefficient i increasing. The sampling data located at the high frequency is not accurate, after rounding off some of the test points at high frequency; we use the low-frequency component ΔR_{th} with more sample data as the object of analysis. Reliable data on low-frequency analysis, the regular pattern is as follows. On one hand, Qp is constant, ΔR_{th} is monotonically increasing with the increase of position coefficient i. With the position increasing to i=16, ΔR_{th} gradually stabilized, increasing gradually smaller, thus the influence of position i on ΔR_{th} is the main factor. On the other hand, the position coefficient i is constant, ΔR_{th} tends to increase monotonically with the increase of Qp. With increasing Qp, ΔR_{th} still tends to be stable, but increase in amplitude is small.

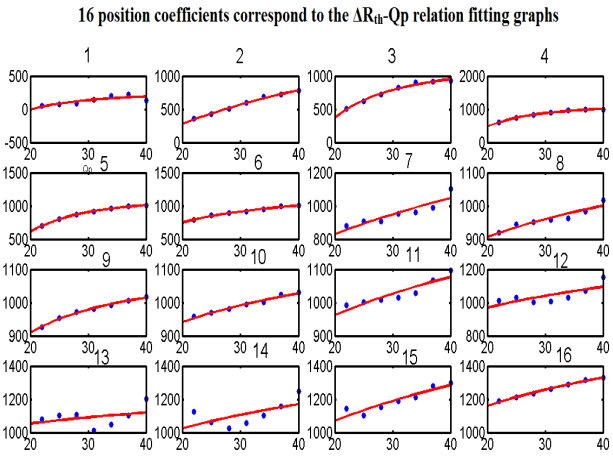

Fig. 5. 16 position coefficients correspond to the ΔR_{th}-Qp relation fitting graphs

According to the above characteristics, we get the two-dimensional function relation of ΔR_{th} and i, Qp, respectively, the analysis is as follows.

Modeling ΔR_{th} relative to coefficient index. According to the measured data to get ΔR_{th} on the position coefficient i of the two-dimensional scatter diagram, as shown in Fig. 6. From the relation of ΔR_{th}-i in the figure, it can be seen that ΔR_{th} obeys the distribution characteristic of logarithmic function (or exponential function log) with respect to position coefficient i. Fitting the scatter plot yields the function of ΔR_{th} with respect to i as (8), (9).

$$\Delta R_{th} = \mu \times e^{a_1 \times i} - \eta \times e^{a_2 \times i} \quad \textbf{(8)}$$

Here, a_1=0.03, a_2=-1.1, μ=774.6, η=2000. Or,

$$\Delta R_{th} = \log_{\tau}(i+\theta) \tag{9}$$

Here, τ=1.0023, θ=0.8572.

Mapping ΔR_{th} to quantization parameter. A two-dimensional scatter plot of the ΔR_{th} with respect to the quantization parameter Qp is obtained from the measured data as shown in Fig. 6.

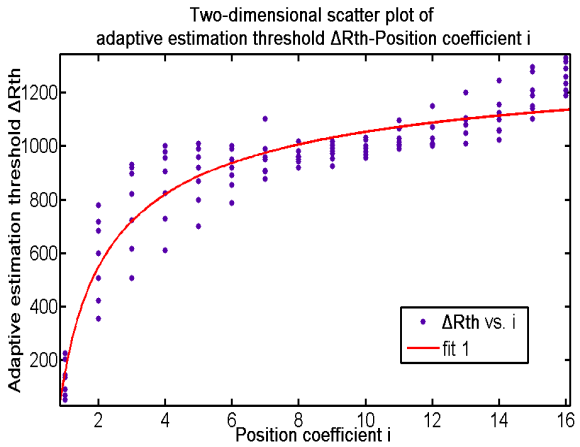

Fig. 6. Two-dimensional scatter plot of the adaptive -estimation threshold ΔR_{th} with respect to the position coefficient i

From the relation of ΔR_{th}-Qp in the figure, ΔR_{th} is monotonically increasing with respect to Qp and becomes stable with increasing Qp. The corresponding function relation is fitted as (10).

$$\Delta R_{th} = b_1 \times (Qp - b_2)^2 + b_3 \tag{10}$$

Here, b_1 = -1.21, b_2 =36.78, b_3 =136.4.

ΔR_{th} modeling. According to analyzing 1, 2, the two-dimensional function ΔR_{th} about i and Qp is added, and the three-dimensional function relation of ΔR_{th} = f (Qp) + f (i) is obtained as formula (11), (12).

$$\Delta R_{th}(Qp, i) = \mu \times e^{a_1 \times i} - \eta \times e^{a_2 \times i} + b_1 \times (Qp - b_2)^2 + b_3 \tag{11}$$

Here, a_1=0.03, a_2=-1.1, μ=774.6, η=2000, b_1 = -1.21, b_2 =36.78, b_3 =136.4.

Or,
$$\Delta R_{th}(Qp, i) = \log_{\tau}(i+\theta) + b_1 \times (Qp - b_2)^2 + b_3 \tag{12}$$

Here, τ=1.0023, θ=0.8572, b_1 = -1.21, b_2 =36.78, b_3 =136.4.

From the data distribution in Fig. 5, we can see that, with the increase of position coefficient i, the data of i sampling point is reduced. After partial test points of i at the high frequency position discarding, the data is filled in the blank position using the model. Fitting 3D surface with ΔR_{th} = f (Qp, i), filling high frequency position ΔR_{th} (blue is the part filled according to the formula). Finetune the threshold after correcting the model data, as shown in Table 1 and Fig. 7.

Table 1. Corrected model ΔR_{th} adaptive threshold

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
22	4.81	473	646	721	764	797	827	857	887	918	950	983	1016	1051	1087	**1124**
25	101	570	743	818	861	893	924	953	983	1014	1046	**1079**	**1113**	**1148**	**1184**	**1221**
28	176	644	817	892	935	968	998	1028	1058	1089	1121	**1154**	**1187**	**1222**	**1258**	**1295**
31	229	697	870	945	988	1021	1051	1081	1111	**1142**	**1174**	**1206**	**1240**	**1275**	**1311**	**1348**
34	260	728	901	976	1019	1052	1082	**1112**	**1142**	**1173**	**1205**	**1237**	**1271**	**1306**	**1342**	**1379**
37	269	737	910	985	1028	1061	1091	**1121**	**1151**	**1182**	**1214**	**1247**	**1281**	**1315**	**1351**	**1388**
40	256	725	898	973	1016	**1049**	**1079**	**1108**	**1139**	**1170**	**1201**	**1234**	**1268**	**1303**	**1339**	**1376**

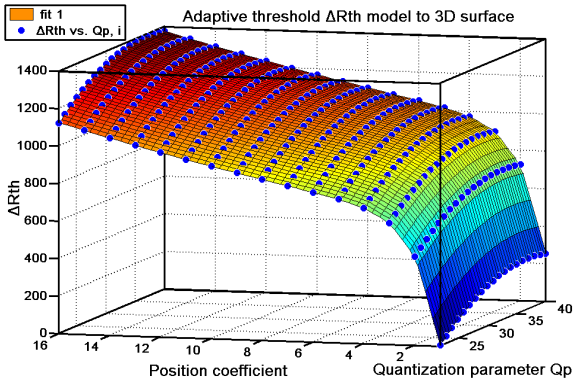

Fig. 7. Fitting adaptive threshold model to 3D surface

3.3 Model Building and Parameter Selection

In [10], offset δ_k is constructed as a function of the quantization parameter Qp and DCT coefficient distribution parameter Λ, as shown in equation (5). Based on this study, we add $\psi_{\Delta R}$ to study the relationship between $\psi_{\Delta R}$ and ideal dead-zone offset, and use $\psi_{\Delta R}$ to mediate the quantization off -set in [10] thereby constructing a deadzone offset model.

In [10], original migration model is δ_k = f (QP, Λ_i), and the adjustment amount ω = f ($\psi_{\Delta R}$) is introduced. Let δ be the new model offset.

$$\delta = (1+\omega)\cdot\delta_k \tag{13}$$

The quantization is realized through dynamic trimming of original offset δ_k by coefficient 1 + ω.

The adjustment parameter ω is related to $\psi_{\Delta R}$, and the scatter plot of δ_k / δ -$\psi_{\Delta R}$ is plotted. δ is the best offset δ_{opt} when the final quantization result of SDQ is 1, 0 in section 2.2. The corresponding relationship and the scatter fit are shown in Fig. 7. In Fig. 8, δ_k / δ and $\psi_{\Delta R}$ satisfy the characteristics of bounded, odd function, positive and negative symmetrical intervals. With the increase or decrease of $\psi_{\Delta R}$, δ_k / δ tend to be stable, which is in accordance with the tangent function.

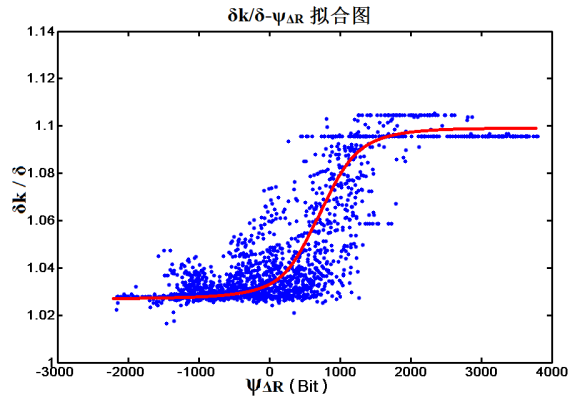

Fig. 8. δ_k / δ-$\psi_{\Delta R}$ scatter fitting

$$\omega = A\arctan(\frac{\psi_{\Delta R}}{B}) \tag{14}$$

From formula (7),

$$\omega = A\arctan(\frac{\Delta R - \Delta R_{th}}{B}) \tag{15}$$

ΔR_{th} is a two-dimensional function of QP and i, such as the formula (11), (12).

$$\psi_{\Delta R} = \Delta R - \Delta R_{th}(Qp,i) \tag{16}$$

A, B respectively adjust the amplitude and slope of ω. (15) into (13), final model such as (17).

$$\delta = (1 + A\arctan(\frac{\Delta R - \Delta R_{th}}{B})) \cdot \delta_k \tag{17}$$

The new model needs to determine optimal combination of parameters A and B. Since ω is a fine adjustment of offset δ, it is assumed that δ is fluctuated from 0.7 to 1.2 times δ_k.

$$\delta = (0.7 \sim 1.2) \cdot \delta_k \tag{18}$$

According to Eq.(17), Eq. (18).

$$0.7 < 1 + A\arctan(\frac{\Delta R - \Delta R_{th}}{B}) < 1.2 \tag{19}$$

Experimentally, the range of A is (-0.1, 0.2), and the range of B is (0.5, 6). As shown in Fig. 9, 73 kinds of A and B combinations were tested and compared with the adjustment amount ω = 0 in the literature [10]. Their rate-distortion performance curves and the three-dimensional surface plot of AB-BD-PSNR were observed. The best combination of (A, B) is (0.3, 6).

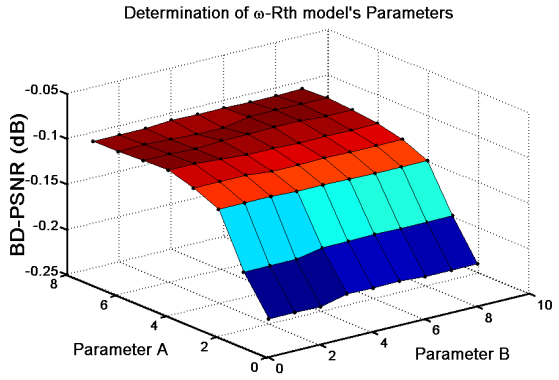

Fig. 9. 3D surface of A-B-BD-PSNR

4 Experimental Results and Analysis

The proposed context-adaptive offset model for deadzone HDQ is applied to H.264 and H.265 standard which is verified by HDQ algorithm. Then the rate-distorted performance of proposed algorithm compares with HDQ algorithm, algorithm in [10] and SDQ algorithm of it. 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 20 frames are tested for all resolution video sequences. The BD-PSNR and BD-RATE are used for performance comparison [15].

For 1080p video sequence, experimental results in Table 2 show that compared with the fixed-offset HDQ algorithm, the proposed algorithm has a performance improvement of 0.0964dB in BD-PSNR, which is equivalent to saving 3.5723% bit rate. Compared with the algorithm in [10], the performance of this algorithm is improved by 0.0188dB, which is equivalent to the saving of 0.6168%. It can be seen that the performance of the algorithm proposed in this paper is better than that of the fixed-offset HDQ algorithm, which is closer to the performance of the optimal SDQ algorithm. This model only increases code rate estimation. The computational time is shown in Table 3. Generally speaking, the additional computation complexity is small. This algorithm can keep the advantages of HDQ algorithm. Fig. 10 shows the performance of SDQ algorithm, [10] algorithm, HDQ and the proposed algorithm.

Table 2. fixed offset HDQ, [10] model, proposed context adaptive self-adaptive offset model and SDQ algorithm performance comparison

Sequences	BD-PSNR(dB)			BD-RATE(%)		Sequences	BD-PSNR(dB)			BD-RATE(%)		
	fixed-offset	Original	proposed	fixed-offset	Original	proposed	fixed-offset	Original	proposed	fixed-offset	Original	proposed
CITY	-0.0909	-0.003	-0.001	2.8722	0.0967	-0.599 cyclists	-0.0685	-0.016	-0.005	2.5625	0.5076	0.0757
HARBOUR	-0.1091	-0.032	-0.004	2.7192	0.7407	0.061 Optis	-0.0833	-0.009	-0.003	2.9077	0.2524	0.0926
husky	-0.4335	-0.139	-0.026	6.6138	2.0178	0.3338 Raven	-0.1384	-0.042	-0.005	4.4667	1.282	0.0894
SOUER	-0.1427	-0.036	-0.02	3.6588	0.8558	-0.5717 vidyo1	-0.0166	0.0071	0.0162	0.2473	-0.582	0.0018
flower-garden	-0.4321	-0.136	-0.027	7.1516	2.1733	0.3796 Kristen AndSara	-0.0363	-0.018	-0.013	1.0851	-0.077	-0.0575
D1 Average	-0.2417	-0.069	-0.016	4.6031	1.1769	-0.0793 720p Average	-0.0686	-0.016	-0.002	2.2539	0.277	0.0404
proposed vs fixed-offset	0.2261			-4.6824		proposed vs fixed-offset	0.0667			-2.2135		
proposed vs Original	0.0537			-1.2562		proposed vs Original	0.0136			-0.2364		
Kimono1	-0.1955	-0.109	-0.055	5.7955	3.515	1.7737 riverbed	-0.1758	-0.056	-0.068	1.6357	1.119	1.59
Tennis	-0.0157	0.004	0.0181	0.7235	-0.153	-0.6925 cactus	-0.0758	0.0261	0.0069	0.6794	-0.8782	-0.2515
sunflower	-0.24	-0.092	-0.003	3.4539	1.9279	-0.3506 rush_hour	-0.101	-0.022	-0.004	1.4429	0.9364	-0.1461
pedestrian_area	-0.0495	-0.015	-0.036	1.8762	0.3772	0.7285 1080pAverage	-0.1219	-0.053	-0.028	2.2296	0.9777	0.3788
proposed vs fixed-offset	0.094			-1.8508		proposed vs Original	0.025			-0.5989		

Table 3. 1080p sequences comparison of time saving for encoding 20 frames

Sequences	SDQ	Original	HDQ	Proposed
Kimono1	2.845	2.645	2.604	2.598
Tennis	2.567	2.525	2.442	2.295
sunflower	1.842	1.861	1.852	1.712
riverbed	4.808	4.785	4.115	3.921
rush_hour	2.321	2.227	2.181	2.218
Cactus	2.328	2.321	2.181	2.217

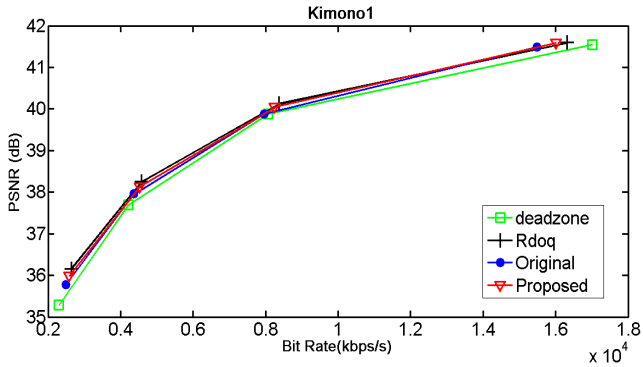

Fig. 10. Performance comparison of four algorithms

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 modelling, this paper proposes a content-adaptive deadzone quantizer to minimize the rate distortion performance difference between the deazone 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 0.08836dB PSNR increment and 3.097% bit rate saving in 1080p sequences. In

addition, this work, in comparison with the SDQ, achieves less than 0.03921dB PSNR loss and 1.51% bit rate increment.

In this paper, we consider correlation between coefficients, therefore, calculating speed of some high-resolution video sequences is not high. The algorithm needs to be further optimized and upgrade in terms of computational complexity. In future, we will focus on realization of algorithm speed, simplify calculation of program, speeding up encoding speed to improve coding efficiency.

Acknowledgements

This work was supported by the Zhejiang NSF LY12F01011.

Reference

[1] G.J. Sullivan, Efficient scalar quantization of exponential and laplacian random variables, Applied IEEE Transactions on Information Theory 42-43(1996) 1365-1374.

[2] G.J. Sullivan, Adaptive quantization encoding technique using an equal expected-value rule, in: Proc. 2005 the 14th Joint Video Team Meeting, 2005.

[3] G.J. Sullivan, S. Sun, On deadzone plus uniform threshold scalar quantization, in: Proc. 2005 the Visual Communication and Image Processing, 2005.

[4] M. Karczewicz, Y. Ye, I. Chong, Rate distortion optimized quantization, Cognition 16(6)(2008) 12-13.

[5] J. Wen, M. Luttrell, J. Villasenor, Trellis-based R-D optimal quantization in H.263+, in: Proc. Applied IEEE Transactions on Image Processing, 2000.

[6] L. Liu, A.M. Yourapis, Rate distortion optimized quantization in the JM reference software, Cognition 16(27) (2008)(24-29).

[7] T.-Y. Huang, H.H. Chen, Efficient quantization based on rate-distortion optimization for video coding, Applied IEEE Transactions on Circuits & Systems for Video Technology, Cognition 26-27(2016) 1099-1106.

[8] H. Lee, S. Yang, Y. Park, B. Jeon, Fast quantization method with simplified rate-distortion optimized quantization for HEVC encoder, Applied IEEE Transactions on Circuits & Systems for Video Technology 10-11(2015) 1-3.

[9] T.-Y. Huang, C.-K. Kao, H.H. Chen, Acceleration of rate-distortion optimized quantization for H.264/AVC, in: Proc. 2013 the IEEE International Symposium on Circuits and Systems, 2013.

[10] H. Yin, H. Wang, X. Wang, Z. Xia, Coefficient-wise deadzone hard-decision quantizer with adaptive rounding offset model, in: Proc. 2016 Data Compression Conference, 2016.

[11] J. Sun, Y. Duan, J. Li, J. Liu, Z. Guo, Rate-distortion analysis of dead-zone plus uniform threshold scalar quantization and its application–part I: fundamental theory, Applied IEEE Transactions on Image Processing 22-23(2013) 202-214.

[12] E.H Yang, S.Y. Shen, Distortion program-size complexity with respect to a fidelity criterion and rate distortion function, Applied IEEE Transactions on Information Theory 39-41(1993) 288-292.

[13] J. He, F. Yang, High-speed implementation of rate-distortion optimized quantization for H.264/AVC, Applied Signal, Applied Image and Video Processing 9-11(2015) 652-661.

[14] J.T. Wen, M. Xiao, J. Chen, P. Tao, C. Wang, Fast rate distortion optimized quantization for H.264/AVC, in: Proc. 2010 the Data Compression Conference, 2010.

[15] G. Bjontegaard, Calculation of average PSNR differences between RD-Curves, Cognition 6(33)(2001) 2-4.