Adaptive Quantization Parameter Cascading in HEVC Hierarchical Coding

Tiesong Zhao, *Member, IEEE*, Zhou Wang, *Fellow, IEEE*, and Chang Wen Chen, *Fellow, IEEE*

***Abstract*— The state-of-the-art High Efficiency Video Coding (HEVC) standard adopts a hierarchical coding structure to improve its coding efficiency. This allows for the quantization parameter cascading (QPC) scheme that assigns quantization parameters (Qps) to different hierarchical layers in order to further improve the rate-distortion (RD) performance. However, only static QPC schemes have been suggested in HEVC test model, which are unable to fully explore the potentials of QPC. In this paper, we propose an adaptive QPC scheme for an HEVC hierarchical structure to code natural video sequences characterized by diversified textures, motions, and encoder configurations. We formulate the adaptive QPC scheme as a nonlinear programming problem and solve it in a scientifically sound way with a manageable low computational overhead. The proposed model addresses a generic Qp assignment problem of video coding. Therefore, it also applies to group-of-picture-level, frame-level and coding unit-level Qp assignments. Comprehensive experiments have demonstrated that the proposed QPC scheme is able to adapt quickly to different video contents and coding configurations while achieving noticeable RD performance enhancement over all static and adaptive QPC schemes under comparison as well as HEVC default frame-level rate control. We have also made valuable observations on the distributions of adaptive QPC sets in the videos of different types of contents, which provide useful insights on how to further improve static QPC schemes.**

***Index Terms*— High efficiency video coding (HEVC), quantization parameter (Qp) cascading, rate-distortion optimization (RDO), hierarchical-B-pictures (HBP), group-of-pictures (GOP), video compression.**

I. Introduction

AIMING at achieving approximately a 50% bit rate reduction for equal quality over H.264 standard, High Efficiency Video Coding (HEVC) [1], [2] adopted a series of new technologies [3], including new prediction structure, larger Coding Tree Units (CTUs), larger size for integer-based Discrete Cosine Transform (DCT) approximations, more intra-prediction modes, merge mode, improved de-blocking filter, etc. To further improve the Rate-Distortion (RD) performance, researchers have been developing new RD Optimization (RDO) approaches that optimize multiple levels of HEVC encoder. These efforts include Coding Unit (CU) level Quantization parameter (Qp) determination [4], [5], Lagrange Multiplier (LM) adaption [6], [7], perceptual LM adaption [8], [9], Structural SIMilarity (SSIM)-inspired [10] LM adaption [11], [12], low complexity RDO [13], [14], and so on. Despite of these great efforts (most of them are at CU-level), the optimization in prediction structure level has not been widely studied and thus the potentials of HEVC improvements have not been fully exploited.

HEVC supports several types of inter-frame prediction structures, including Random Access (RA), Low Delay (LD) and Low Delay P (LP), to fulfill diversified requirements. Each structure is implemented with a set of hierarchical P or B pictures within a Group-of-Picture (GOP) [15], where a picture in a higher layer can be predicted from one or more pictures from its lower layers. Such a prediction order implies that, compared with the higher layers, a lower layer may have a greater impact on the overall RD performance. Therefore, it is feasible to apply a different Qp in each layer in order to maximize the overall RD performance. This strategy yields the so-called Quantization Parameter Cascading (QPC) approach. Such an idea remains high level without full implementations in specific encoders. In HEVC standard description, it is unclear how to adaptively determine an optimal Qp set by providing an average Qp. Since QPC approaches have been developed for legacy video coding standards, it is beneficial to review several existing QPC approaches and discuss the needs for developing a new adaptive QPC for HEVC hierarchical encoding schemes. These existing schemes can be categorized into three groups.

1) Static QPC Methods: A static QPC method is essentially an heuristic method derived by empirical data, in which the optimal Qp set is selected through extensive experiment. The first and most well-known static QPC model was proposed in JVT-P014 [15] as:

$$\mathrm{Qp}_k = \mathrm{Qp}_0 + \Delta\mathrm{Qp}_{base} + (k-1), k > 0 \tag{1}$$

where Qp_k and Qp_0 denote the Qp values of layer k and layer 0 (*i.e.*, the lowermost layer), respectively; and $\Delta\mathrm{Qp}_{base} = 4$ is the offset base. This model was initially designed for H.264 hierarchical B coding and has shown significant improvement

Manuscript received August 26, 2015; revised December 15, 2015 and April 2, 2016; accepted April 8, 2016. Date of publication April 20, 2016; date of current version May 11, 2016. This work was supported in part by the U.S. National Science Fundation (NSF) under Grant ECCS-1405594, in part by the National Sciences and Engineering Research Council (NSERC) of Canada under the Discovery, Strategic, Idea-to-Innovation and Steacie Award programs, in part by the NSF China under Grant 61571129 and in part by the Cross-strait joint fund of NSF China under Grant U1405251. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yui-Lam Chan.

T. Zhao is now with the College of Physics and Information Engineering, Fuzhou University, Fuzhou 350116, China. He was with the the Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: t.zhao@fzu.edu.cn).

Z. Wang is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: z.wang@ece.uwaterloo.ca).

C. W. Chen is with the Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: chencw@buffalo.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2556941

in the overall coding efficiency [16]. Since then, it has attracted substantial attention in the video coding community. Variations of this model have been employed as the default QPC schemes for both Scalable Video Coding (SVC) [17] and HEVC test model (HM) [18]. In these two scenarios, the values of $\Delta \mathrm{Qp}_{base}$ were chosen to be 3 and 1, respectively. Another variation of JVT-P014 was proposed by Jing *et al.* [19] also for SVC. In their work, Eq. (1) with $\Delta \mathrm{Qp}_{base} = 3$ was employed in the reference software encoder to calculate the optimal Qp set except the Qp of the highest hierarchical layer, which was set as the average Qp of all other layers.

2) Adaptive QPC Methods: In an adaptive scenario, the optimal Qp set of different GOPs can be dynamically updated based on real-time statistics of the coding features. In [20], Wan *et al.* developed a content adaptive QPC scheme for hierarchical P picture structure with H.264/AVC baseline profile. In this case, the Qp of the highest hierarchical layer was determined according to the desired target bit rate. Then, Qps of all other layers were designated based on the above Qp and an offset that was indirectly obtained by summarizing prediction modes of all layers. An adaptive QPC scheme for SVC was proposed by Li *et al.* [21], where a linear factor was introduced to represent the error propagation due to quantization in each layer. Based on RD curve emulations, they have designed a scheme to dynamically obtain the Qp of the lowermost layer, Qp_0, and the optimal Qp offsets of all other layers (i.e., $\mathrm{Qp}_k - \mathrm{Qp}_0$ for all $k > 0$) in order to obtain the optimal Qp set. In [22], Li *et al.* presented an improved model over the original one [21] by minimizing the overall distortion of the entire GOP, subject to an average bit rate of all frames within a GOP. With this model, it is impractical to obtain an analytical solution of global optimum due to its mathematical complexity. Instead, the Qp_0 was updated in a similar way as in [21] and the optimal Qp offsets of other layers were obtained empirically and stored in several lookup tables before coding a video sequence. It was shown that good improvements were achieved for both hierarchical B and hierarchical P pictures.

3) QPC Schemes Integrated With Other Coding Components: Various QPC schemes have also been integrated into other video coding algorithms [23]–[27] to further improve the RD performance. QPC was utilized to improve the multi-point video conferencing system [23], [24], where the temporally scaled video codec with QPC was shown to be comparable to the non-scaled version in terms of RD performance. In [25] and [26], the QPC scheme was combined with GOP-based rate control to develop single-pass dependent bit allocation methods for H.264 and SVC, respectively. In [27], QPC was adopted in Qp clipping aiming at a rate control algorithm for HEVC hierarchical structure.

It is clear from the above discussion that an adaptive QPC scheme is preferred over static ones in order to meet various video coding requirements resulting from different video contents as well as encoder configurations (*e.g.*, Qp settings). However, most existing QPC schemes, including both static and adaptive ones, have been developed for H.264 encoders and may not be applicable to HEVC encoder. This is because enormous new features have been introduced in HEVC which may cause possible performance degradation if an existing QPC scheme is directly applied. To address this issue, we propose an approach to adaptively optimize the quantization set of all hierarchical layers. We formulate the adaptive QPC problem as a non-linear programming problem and solve it by utilizing the inherent inter-layer RD dependencies and adaptively updating the RD models for each GOP. The proposed adaptive QPC scheme has been validated by simulations on numerous video sequences with different spatial resolutions, frame rates, prediction structures and Qp settings. These simulation results have demonstrated high efficiency, low complexity and fast adaptation of the proposed QPC scheme. The statistics of adaptive QPC set distributions in videos of different contents can also provide insights on how to further improve static QPC schemes. The proposed model also applies to other Qp assignment problems in video coding including Qp assignments to GOPs, frames and CUs. Therefore, it provides a new vision to address Qp assignment problems in HEVC coding.

The rest of the paper is organized as follows. Sections II and III formulate the optimization problem of adaptive QPC and solve the problem with non-linear programming, respectively. Extensive experiments are studied in Section IV with comparisons against recently developed schemes. Finally, Section V concludes the paper with a summary and future outlooks.

II. The Adaptive QPC Problem in Hierarchical Video Coding

We formulate the QPC problem as follows. Assume in a GOP with L hierarchical layers, the average quantization stepsize (Q) is set as $\bar{Q}$. To assign Q to each hierarchical layer, one simple way is to use the same $\bar{Q}$ for each layer, and another way is to use different values Q_l, $l = 0, 1, 2, \ldots, L-1$, for different layers. However, for layer l, replacing $\bar{Q}$ with Q_l may lead to increments in bit rate and the total distortion of the whole GOP. We hereinafter denote these by $\Delta D_{tot,l}$ and $\Delta R_{tot,l}$, respectively. The goal of adaptive QPC optimization is to find an optimal set of Q_l, $l = 0, 1, 2, \ldots, L-1$, that minimizes the overall distortion of a GOP without increasing the total number of bits. This optimization is given by:

$$\min_{\mathbf{Q}} \left\{ \sum_{l=0}^{L-1} \Delta D_{tot,l} \right\}, \quad s.t. \ \sum_{l=0}^{L-1} \Delta R_{tot,l} \leq 0, \tag{2}$$

where $\mathbf{Q}$ denotes the optimal set of $Q_l, l = 0, 1, 2, \ldots, L-1$. $\Delta D_{tot,l}$ and $\Delta R_{tot,l}$ can be affected by several related factors including $\bar{Q}$, Q_l, Distortion-Q (D-Q) and Rate-Q (R-Q) models.

A. Inter-Layer RD Dependencies

An example of the RA hierarchical structure is presented in Fig. 1, where only 1 reference picture is shown in each direction. The arrows with dashed lines show predictions between hierarchical layers in each GOP. It can be easily observed that the RD performance of a frame/slice may be

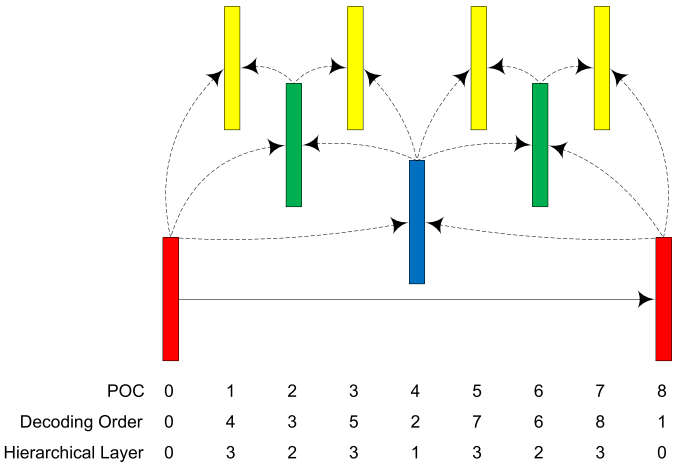


Fig. 1. An example of an RA prediction structure with a GOP size of 8.

influenced by its reference frames in a lower layer. This RD dependency was studied in [28] for HEVC default RA setting. An approximately linear inter-layer D dependency was observed as

$$\frac{\Delta D_{cur}}{\Delta D_{ref}} = \delta, \tag{3}$$

where D_{cur} and D_{ref} denote the distortions (in terms of Mean Squared Error, or MSE) of a frame/slice and its reference, respectively; δ represents the intensity of error propagation. On the other hand, low dependency between the corresponding bit rates, R_{cur} and R_{ref}, was observed.

Brief explanations of the above dependencies are as follows. In a hierarchical structure, the CUs are highly probable to be predicted by inter frames. Therefore, the reconstruction errors of a reference frame are readily propagated to higher layers. In such a case, the values of D_{cur} and D_{ref} are positively correlated, which leads to a positive δ. Usually, $\delta \in [0, 1]$. $\delta = 0$ indicates that there is no inter-layer error propagation, *i.e.*, the distortion of a lower layer shall not result in any distortion in the higher layers; At the other extreme, $\delta = 1$ indicates full error propagation. In this work, we generally tune δ using video sequences provided by HEVC Common Test Condition (CTC) [29] and obtain $\delta = 0.5$. For the inter-layer R dependency, the coding bit rates are mostly determined by the Qp, motion search and entropy coding and they are barely influenced by error propagation. Therefore, we assume a zero R dependency.

In real-life HEVC encoders, multiple reference pictures are employed, resulting in more than one D_{ref} values corresponding to one D_{cur}. However, according to our observation, there still exists a high probability that the nearest reference in each direction is selected. This conclusion can be validated with statistics of HEVC encoded videos. In Table I, we present the percentages of using the nearest references in HEVC RA. The default RA setting is used to test 12 video sequences and 4 Qp settings for each sequence. From the table, it is probable that the nearest reference is employed to predict the current frame/slice, especially for those cases where a larger Qp setting is used. Therefore it is reasonable to utilize the nearest reference only in the following derivation process,

TABLE I
PERCENTAGES OF USING THE NEAREST REFERENCES IN HEVC RA

Video	Qp=22	Qp=27	Qp=32	Qp=37
RaceHorses 832×480	83.6	85.3	86.4	87.5
BQMall 832×480	84.0	86.0	88.1	89.1
PartyScene 832×480	85.7	86.6	88.8	90.4
BasketballDrill 832×480	81.1	82.4	82.9	84.5
RaceHorses 416×240	86.9	87.8	88.5	88.4
BQSquare 416×240	80.4	81.1	83.7	89.5
BlowingBubbles 416×240	87.3	88.7	89.8	90.3
BasketballPass 416×240	87.0	88.2	90.3	90.3
BaskeballDrillText 832×480	81.8	82.0	83.0	83.7
ChinaSpeed 1024×768	82.2	82.9	84.3	84.6
SlideEditing 1280×720	78.8	79.9	84.7	87.9
SlideShow 1280×720	88.4	88.9	89.4	89.5
Average	83.9	85.0	86.7	88.0

which is similar to that in [28]. The coding performance of our conclusion will be examined by HEVC default multiple reference settings in Section V.

Due to the inter-layer D dependency presented above, the average distortion of layer l, ΔD_l, has an impact on the incremental distortion $\Delta D_{tot,l}$. Following the derivations in [28] and [30], we obtain

$$\Delta D_{tot,l} = \alpha_l \Delta D_l, \tag{4}$$

where

$$\alpha_l = \begin{cases} \dfrac{1}{1-\delta}(1+2\delta)^{L-1} & \text{if } l = 0, \\ (1+2\delta)^{L-1-l} & \text{otherwise.} \end{cases} \tag{5}$$

The inter-layer R dependency can be approximated as 0, therefore, the incremental bit rates $\Delta R_{tot,l}$ is obtained as the sum of ΔR_l of all frames in the same layer:

$$\Delta R_{tot,l} = \beta_l \Delta R_l, \tag{6}$$

where

$$\beta_l = \begin{cases} 1 & \text{if } l = 0, \\ 2^{l-1} & \text{otherwise.} \end{cases} \tag{7}$$

B. Adaptive D-Q and R-Q Models

In the optimization in Eq. (2), the changes in D_l, R_l result from changing the parameter Q. Hence, we need to determine appropriate D-Q and R-Q models for HEVC for the purpose of adaptive QPC. In recent years, researchers have published numerous D-Q and R-Q models for video coding to improve RDO and associated techniques. Several popular D-Q and R-Q models are summarized as follows, which include linear D-Q model [31], quadratic D-Q model [32], Cauchy distribution-based D-Q model [33], linear R-Q model [34] and two quadratic R-Q models [35], [36].

Considering that these models were developed for H.264 and earlier standards, it is not clear whether they work well for a large range of Qps with an HEVC encoder. To compare them, we test seven 832×480 benchmark videos with HEVC RA and the Qp from 22 to 42 to obtain a series of D, R, Q and MAD values. Firstly, we test the model accuracy with the average D, R, Q and MAD values of all hierarchical layers.

TABLE II
COMPARISON I OF DIFFERENT D-Q AND R-Q MODELS

Video	D-Q [31]		D-Q [32]		D-Q [33]		R-Q [34]		R-Q [35]		R-Q [36]	
	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE
BasketBall	0.9908	3.25	0.9512	12.93	0.9895	3.85	0.7247	0.0539	0.8205	0.0450	0.8169	0.0451
BQmall	0.9869	5.77	0.9589	15.22	0.9865	5.47	0.6060	0.0950	0.7261	0.0826	0.7050	0.0847
Flowervas	0.9670	2.31	0.9369	4.61	0.9668	2.29	0.5712	0.0216	0.7031	0.0188	0.6832	0.0192
Keiba	0.9409	10.52	0.9089	17.52	0.9397	10.63	0.6203	0.1262	0.7807	0.1010	0.7562	0.1052
Mobisode2	0.5215	6.68	0.4937	7.47	0.5240	6.83	0.5040	0.0397	0.4459	0.0412	0.5648	0.0379
PlayScene	0.9786	15.95	0.9452	32.73	0.9743	17.69	0.6157	0.1799	0.7991	0.1379	0.7660	0.1468
RaceHorses	0.9911	6.16	0.9513	23.83	0.9881	8.94	0.5383	0.1950	0.8463	0.1240	0.8115	0.1352
Average	0.9109	7.23	0.8780	16.33	0.9098	7.96	0.5972	0.1016	0.7317	0.0786	0.7291	0.0820

TABLE III
COMPARISON II OF DIFFERENT D-Q AND R-Q MODELS

Video	D-Q [31]		D-Q [32]		D-Q [33]		R-Q [34]		R-Q [35]		R-Q [36]	
	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE
BasketBall	0.9960	2.15	0.9687	10.73	0.9952	2.52	0.9320	0.0244	0.9778	0.0137	0.9802	0.0135
BQmall	0.9935	4.75	0.9793	11.89	0.9946	3.32	0.9281	0.0361	0.9476	0.0313	0.9553	0.0292
Flowervas	0.9816	1.88	0.9692	3.58	0.9828	1.68	0.9082	0.0064	0.9266	0.0056	0.9545	0.0026
Keiba	0.9478	9.36	0.9267	14.66	0.9471	9.33	0.9119	0.0587	0.9346	0.0557	0.9545	0.0438
Mobisode2	0.5377	6.44	0.5176	7.12	0.5385	6.62	0.7997	0.0239	0.5572	0.0365	0.8130	0.0236
PlayScene	0.9860	14.73	0.9702	24.51	0.9860	12.40	0.8920	0.0920	0.9515	0.0553	0.9668	0.0470
RaceHorses	0.9941	5.82	0.9676	18.94	0.9918	7.20	0.9004	0.0869	0.9939	0.0179	0.9938	0.0180
Average	0.9195	6.45	0.8999	13.06	0.9194	6.15	0.8960	0.0469	0.8984	0.0309	0.9454	0.0254

For each video and Qp, the average values are fitted with different D-Q and R-Q models; then the fitted $\{\hat{D}_{Qp}|Qp = 22\ldots42\}$ and $\{\hat{R}_{Qp}|Qp = 22\ldots42\}$ are compared with the "actual" $\{D_{Qp}|Qp = 22\ldots42\}$ and $\{R_{Qp}|Qp = 22\ldots42\}$ values, with the corresponding Pearson Linear Correlation Coefficient (PLCC) and Root Mean Square Error (RMSE) values shown in Table II. Secondly, we follow the above process to test the model accuracy over each hierarchical layer; then the PLCC and RMSE values of all layers are averaged and presented in Table III. For computational convenience, the D-Q model of [33] is calculated in logarithmic domain to minimize the squared error.

From Tables II and III, the D-Q models [31], [33] and R-Q model [36] perform relatively better for these benchmark videos. For a trade-off between accuracy and complexity, we adopt the linear D-Q model [31] and quadratic R-Q model [36] in this research. They can be respectively represented by

$$D = \gamma\, Q, \tag{8}$$

and

$$R = \frac{\mu m}{Q^2} + \frac{\nu m}{Q} + \phi, \tag{9}$$

where γ is the D-Q model parameter; μ and ν are the second and first order coefficients, respectively, and ϕ represents the header bits.

Another observation from Tables II and III is, accurate estimates of D-Q and R-Q models in different hierarchical layers lead to both higher correlations and lower prediction errors. This is especially true for the R-Q model. The reason may be due to significant variations of the header bits, especially the motion vector bits, in different hierarchical layers which are in turn due to temporal distances to the reference frames. Based on this, it is desired to predict more accurate γ_l, m_l, μ_l and ν_l for all hierarchical layers. To make the QPC dynamically adapted to different video sequences with various textures, motions and encoder configurations, we progressively update the layer parameters in the coding procedure,

$$\begin{cases} \hat{\gamma}_l(n+1) = \hat{\gamma}_l(n)\cdot(1-w_d) + \gamma_l(n)\cdot w_d, \\ \hat{m}_l(n+1) = \hat{m}_l(n)\cdot(1-w_m) + m_l(n)\cdot w_m, \\ \hat{\mu}_l(n+1) = \hat{\mu}_l(n)\cdot(1-w_r) + \mu_l(n)\cdot w_r, \\ \hat{\nu}_l(n+1) = \hat{\nu}_l(n)\cdot(1-w_r) + \nu_l(n)\cdot w_r, \end{cases} \tag{10}$$

where n indicates the GOP index; $\hat{\gamma}_l$, $\hat{m}_l$, $\hat{\mu}_l$ and $\hat{\nu}_l$ represent the predicted value of parameters. w_d, w_r and w_m denote the updating coefficients of D-Q, R-Q and MAD models, respectively. In this work, they are obtained by tuning γ_l, m_l, μ_l and ν_l values of Class C & D CTC sequences [29]. We set $w_d = 0.87$, $w_r = 0.25$ and $w_m = 0.84$, respectively.

C. The Adaptive QPC Problem

By applying the above R-Q and D-Q models to all hierarchical layers, we obtain

$$\Delta D_l = \gamma_l \Delta Q_l, \tag{11}$$

and

$$\Delta R_l = -\left\{\frac{2\mu_l m_l}{Q_l^3} + \frac{\nu_l m_l}{Q_l^2}\right\}\Delta Q_l. \tag{12}$$

Substitute Eqs. (4), (6), (11), (12) into Eq. (2), we formulate the adaptive QPC as the following constrained

optimization problem:

$$\min_{\mathbf{Q}} \left\{ \sum_{l=0}^{L-1} \alpha_l \gamma_l \Delta Q_l \right\},$$
$$s.t. \ -\sum_{l=0}^{L-1} \left\{ \frac{2\beta_l \mu_l m_l}{(\bar{Q}+\Delta Q_l)^3} + \frac{\beta_l \nu_l m_l}{(\bar{Q}+\Delta Q_l)^2} \right\} \Delta Q_l \le 0. \tag{13}$$

We can then obtain the optimal Q set for all hierarchical layers by solving this problem.

Eq. (13) applies to several Qp assignment problems in HEVC encoders. These problems include Qp assignments from video sequence level to GOPs, from GOPs to hierarchical layers, from layers to frames, and from frames to CUs. In each step, we can formulate the Qp assignment problem as Eq. (13), where $\bar{Q}$ represents the average Q of a video sequence, a GOP, a layer or a frame and L represents the number of GOPs, layers, frames or CUs. The only difference is, we need to explore the inter-GOP, inter-layer, inter-frame or inter-CU Rd dependencies that may lead to different α_l and β_l values. In this sense, the proposed model provides a new solution to the Qp assignment problems in HEVC. In the following Section IV.D, the application of Eq. (13) in inter-GOP Qp assignment will be presented for LD and LP encoders, in which comprehensive experiments demonstrate the efficiency of the proposed model.

III. The Proposed Solution

In Eq. (13), let $a_l = -\alpha_l \gamma_l$, $b_l = 2\beta_l \mu_l m_l$, $c_l = \beta_l \nu_l m_l$, $d = \bar{Q}$ and $x_l = \Delta Q_l$, we have

$$\max_{\mathbf{x}} \left\{ \sum_{l=0}^{L-1} a_l x_l \right\},$$
$$s.t. \ \sum_{l=0}^{L-1} \left\{ \frac{b_l}{(x_l+d)^3} + \frac{c_l}{(x_l+d)^2} \right\} x_l \ge 0, \tag{14}$$

where $\mathbf{x}$ denotes the solution vector $[x_0; x_1; \ldots; x_{L-1}]$. Eq. (13) can be considered as a non-linear programming problem.

A. The Determination of Qps

By introducing the Lagrange multiplier Λ in the non-linear programming problem, we obtain

$$\mathbb{L} = \sum_{l=0}^{L-1} a_l x_l + \Lambda \sum_{l=0}^{L-1} \left\{ \frac{b_l}{(x_l+d)^3} + \frac{c_l}{(x_l+d)^2} \right\} x_l. \tag{15}$$

Let vector $\mathbf{x_E} = [\mathbf{x}; \Lambda]$ and set $\frac{\partial \mathbb{L}}{\partial \mathbf{x_E}} = 0$, we have

$$\mathbf{f}(\mathbf{x_E}) = [f_0(\mathbf{x_E}); f_1(\mathbf{x_E}); \ldots; f_{L-1}(\mathbf{x_E}); f_L(\mathbf{x_E})] = 0, \tag{16}$$

where

$$f_0(\mathbf{x_E}) = \sum_{l=0}^{L-1} \left\{ \frac{b_l}{(x_l+d)^3} + \frac{c_l}{(x_l+d)^2} \right\} x_l, \tag{17}$$
$$f_{l+1}(\mathbf{x_E}) = a_l (x_l+d)^4 - \Lambda \left[b_l (2x_l - d) + c_l \left(x_l^2 - d^2 \right) \right], \tag{18}$$

for $l = 0, 1, 2, \ldots, L-1$.

Eq. (16) can be solved with the Newton-Raphson method. Let

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_0}{\partial x_0} & \frac{\partial f_0}{\partial x_1} & \cdots & \frac{\partial f_0}{\partial x_{L-1}} & \frac{\partial f_0}{\partial \Lambda} \\ \frac{\partial f_1}{\partial x_0} & \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_{L-1}} & \frac{\partial f_1}{\partial \Lambda} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial f_{L-1}}{\partial x_0} & \frac{\partial f_{L-1}}{\partial x_1} & \cdots & \frac{\partial f_{L-1}}{\partial x_{L-1}} & \frac{\partial f_{L-1}}{\partial \Lambda} \\ \frac{\partial f_L}{\partial x_0} & \frac{\partial f_L}{\partial x_1} & \cdots & \frac{\partial f_L}{\partial x_{L-1}} & \frac{\partial f_L}{\partial \Lambda} \end{bmatrix}$$
$$= \begin{bmatrix} \mathrm{A}_0 & \mathrm{A}_1 & \cdots & \mathrm{A}_{L-1} & 0 \\ B_0 & & & & \Gamma_0 \\ & B_1 & & & \Gamma_1 \\ & & \ddots & & \vdots \\ & & & B_{L-1} & \Gamma_{L-1} \end{bmatrix}, \tag{19}$$

where

$$\mathrm{A}_l = \frac{\Gamma_l}{(x_l+d)^4},$$
$$\mathrm{B}_l = 4a_l (x_l+d)^3 - 2\Lambda (b_l + c_l x_l),$$
$$\Gamma_l = -b_l (2x_l - d) - c_l \left(x_l^2 - d^2 \right). \tag{20}$$

The vector $\mathbf{x_E}$ can be iteratively obtained by

$$\mathbf{x_E}(t+1) = \mathbf{x_E}(t) - \mathbf{J}^{-1} \mathbf{f}(\mathbf{x_E}(t)), \tag{21}$$

where t denotes the number of iterations.

According to [37], the convergence is quadratic if A_l, B_l and Γ_l are sufficiently smooth and the initial point is close to one of the roots of the equations. In this work, we set $\mathbf{x_E}(0) = \{0; 0; \ldots; 0; 1\}$ and the iteration is stopped when $\|\mathbf{J}^{-1}\mathbf{f}(\mathbf{x_E}(t))\|^2 \le 0.01 \cdot (L+1)$, which yields an average error of ΔQ_l to be approximately less than 0.1. The iteration process usually takes less than 10 times. Note that there exists a very small possibility that $\mathbf{J}^{-1}$ does not exist if

$$\det(\mathrm{J}) = (-1)^{L-1} \sum_{i=0}^{L-1} \left\{ \mathrm{A}_i \Gamma_i \prod_{j=0, j\ne i}^{L-1} \mathrm{B}_j \right\} = 0. \tag{22}$$

In such a case, or in the case that the iteration process does not converge after 100 iterations, we terminate the iteration and use a static Qp scheme, which will be discussed in Section III.B. The computational complexity of the proposed scheme is negligible in an HEVC encoder, which will be discussed later in more details.

After getting the solution $\hat{\mathbf{x}}_{\mathbf{E}} = [\hat{x}_0; \hat{x}_1; \ldots; \hat{x}_{L-1}; \hat{\Lambda}]$, we can determine the optimal set $\hat{Q}_l, l = 0, 1, 2, \ldots, L-1$ by

$$\hat{Q}_l = \bar{Q} + \hat{x}_l. \tag{23}$$

The last step of the proposed scheme is to find the corresponding Qp_l of each $\hat{Q}_l$. Because the quantization step sizes are discontinuous with a limited number, we need to find a quantization step size, $\tilde{Q}_l$, which is closest to $\hat{Q}_l$. According to [38], there exists an approximately exponential relationship between Qp and Q,

$$Q \approx C_1 \cdot C_2^{\mathrm{Qp}}, \tag{24}$$

TABLE IV
COMPARISON OF DIFFERENT METHODS WITH HEVC RA

Video		SVC [17]		Li's [22]		F-RC		Proposed	
		BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR
Class A	*Traffic*	-0.0781	2.32	-0.1205	3.41	-0.0076	0.08	0.0578	-1.71
	PeopleOnStreet	-0.0975	2.20	-0.1394	3.07	-0.0323	0.76	0.0299	-0.66
	Nebuta	-0.0602	1.97	-0.0757	58.31	0.0209	1.44	-0.0114	-1.65
	SteamLocomotive	-0.1042	8.21	-0.1780	14.79	-0.1526	10.52	-0.0027	0.30
	Average	-0.0850	3.67	-0.1284	19.90	-0.0429	3.20	0.0184	-0.93
Class B	*Kimono*	-0.0848	2.83	-0.1344	4.38	-0.1094	3.75	0.0021	0.26
	ParkScene	-0.0711	2.23	-0.1159	3.59	-0.0364	1.19	0.0180	-0.54
	Cactus	-0.0769	3.50	-0.1461	4.59	-0.0036	0.52	0.0182	-1.08
	BQTerrace	-0.0759	6.05	-0.0998	18.38	-0.0094	1.42	0.0023	0.51
	BasketballDrive	-0.0944	4.42	-0.1180	4.28	-0.0063	0.46	0.0057	-0.26
	Average	-0.0806	3.80	-0.1229	7.05	-0.0330	1.47	0.0093	-0.22
Class C	*RaceHorses*	-0.1136	3.06	-0.1440	3.96	0.0712	-1.84	0.0052	-0.14
	BQMall	-0.0196	0.53	-0.0212	0.46	0.0633	-1.67	0.0443	-1.14
	PartyScene	0.0100	-0.24	-0.0128	0.34	0.0936	-2.31	0.0976	-2.32
	BasketballDrill	0.0360	-0.83	0.0606	-1.37	-0.0202	0.46	0.1516	-3.61
	Average	-0.0218	0.63	-0.0293	0.85	0.0520	-1.34	0.0747	-1.80
Class D	*RaceHorses*	-0.0878	1.89	-0.1252	2.96	-0.0468	1.05	0.0306	-0.67
	BQSquare	-0.0200	0.55	-0.0058	0.19	-0.1076	3.03	0.0479	-1.31
	BlowingBubbles	-0.0081	0.20	-0.0204	0.52	0.0291	-0.67	0.0670	-1.61
	BasketballPass	-0.0418	0.87	0.0104	-0.20	0.1630	-3.32	0.0671	-1.37
	Average	-0.0394	0.88	-0.0353	0.87	0.0094	0.02	0.0532	-1.24
Class F	*BaskeballDrillText*	0.0185	-0.39	0.0551	-1.14	-0.0332	0.73	0.1454	-3.21
	ChinaSpeed	-0.0946	1.80	-0.1057	2.08	0.0070	-0.15	0.0896	-1.69
	SlideEditing	-0.0804	0.52	-0.1513	0.98	-1.1757	8.21	-0.0464	0.29
	SlideShow	0.0679	-0.78	0.0468	-0.53	-0.6910	8.96	0.1085	-1.30
	Average	-0.0222	0.29	-0.0388	0.35	-0.4732	4.44	0.0742	-1.48
Average		-0.0513	1.95	-0.0734	5.86	-0.0945	1.55	0.0443	-1.09

where C_1 and C_2 are two constants. Hence, we determine $\tilde{Q}_l$ as the quantization step size that is least different from $\hat{Q}_l$ in logarithmic domain. Qp_l is then set as the corresponding Qp of $\tilde{Q}_l$. Specially, to avoid quality fluctuation between GOPs, we clip the Qps of all layers to be within $[\bar{Q}_p - 3, \bar{Q}_p + 3]$, where $\bar{Q}_p$ is the average Qp of all layers calculated by the static Qp scheme.

The above clipping process may not guarantee a non-positive bit rate increase because the quantization sizes of all layers have been changed slightly. To avoid the undesirable increase in bit rate, we use the above approximation process to obtain $Qp_l, l = 0, 1, 2, \ldots, L-2$ and set

$$\begin{aligned}\mathrm{Qp}_{L-1} &= \left\lfloor \frac{1}{\beta_{L-1}} \left(\bar{Q}_p \cdot \sum_{l=0}^{L-1} \beta_l - \sum_{l=0}^{L-2} \mathrm{Qp}_l \beta_l \right) \right\rfloor \\ &= \left\lfloor \bar{Q}_p + \sum_{l=0}^{L-2} \frac{\beta_l}{\beta_{L-1}} \left(\bar{Q}_p - \mathrm{Qp}_l \right) \right\rfloor, \end{aligned} \tag{25}$$

where β_l represents the number of frames in layer l and $\lfloor \cdot \rfloor$ denotes the floor function to force the Qp value to be an integer.

B. The Overall QPC Scheme

We summarize the entire adaptive QPC scheme as follows.

Step 1. Initialize the basic Qp to be Qp_{init}. Encode the first I frame with the Qp of $\mathrm{Qp}_{init} - 1$. Determine the Qp set of the first hierarchical GOP via

$$\mathrm{Qp}_l = \begin{cases} \mathrm{Qp}_{init} + 1 & \text{if } l = 0, \\ \mathrm{Qp}_{l-1} + 1 & \text{otherwise.} \end{cases} \tag{26}$$

Set $\bar{Q}$ as the average Q values of all frames in the first GOP.

Step 2. Encode the first hierarchical GOP. Initialize the R-Q and D-Q model parameters. Calculate $\gamma_l(0)$ and $m_l(0)$ of each hierarchical layer l based on the coding results of the first GOP. Since the R-Q model is a quadratic function of MAD and Q, we cannot acquire $\mu_l(0)$ and $\nu_l(0)$ for each layer l now. Instead, we assume only one R-Q model for all hierarchical layers of the first GOP, and get the same $\mu_l(0)$ and $\nu_l(0)$ values for all layers. Then, we set $\hat{\gamma}_l(0) = \gamma_l(0)$, $\hat{m}_l(0) = m(0)$, $\hat{\mu}_l(0) = \mu_l(0)$ and $\hat{\nu}_l(0) = \nu_l(0)$.

Step 3. Perform QPC for the next GOP. Apply the D-Q and R-Q models and MAD in Eq. (13) to get the optimal $\hat{Q}_l, l = 0, 1, 2, \ldots, L-1$. Find the optimal Qp set for all hierarchical layers based on approximation and Eq. (25) as described in Section III.A. For the special case when $\det(\mathrm{J}) = 0$ or the iteration process does not converge in the Newton-Raphson method, or when the same Qp are derived for all layers, the Qp cascading scheme of previous GOP is used instead.

Step 4. Encode this GOP with the derived optimal Qp set.

Step 5. Update the D-Q and R-Q models. Calculate $\gamma_l(n)$, $m_l(n)$, $\mu_l(n)$ and $\nu_l(n)$ for each hierarchical layer l by taking coding results of the most recent two GOPs into consideration. If the same Qp set is used in the two GOPs, which indicates that the quadratic R-Q model parameters cannot be estimated for each layer, we generally assume $\mu_l(n)$ and $\nu_l(n)$ are the same for all layers and estimate these two parameters from all

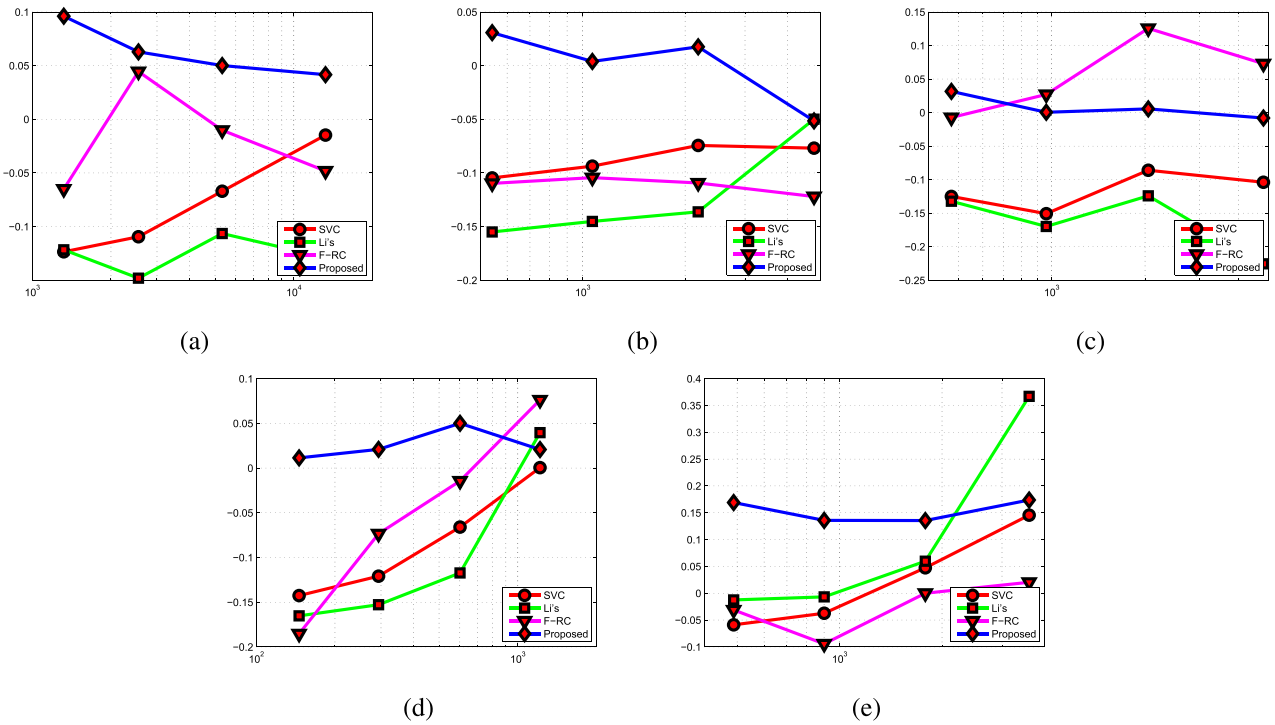


Fig. 2. Examples of Y-PSNR improvements at the same bit rates. Horizontal axis: Bit rate; Vertical axis: Y-PSNR improvement. (a) *Traffic* in Class A. (b) *Kimono* in Class B. (c) *RaceHorses* in Class C. (d) *RaceHorses* in Class D. (e) *BasketballDrillText* in Class F.

TABLE V
COMPUTATIONAL COMPLEXITIES OF NEWTON-RAPHSON METHOD IN TERMS AVENUMITERPERGOP, AVETIMEPERITER (μs) AND TOTITERTIMEPPM

Video		Qp=22	Qp=27	Qp=32	Qp=37
Class C	*RaceHorses*	4.74 / 0.35 / 0.10	5.82 / 0.21 / 0.09	6.79 / 0.19 / 0.11	6.79 / 0.18 / 0.13
	BQMall	6.31 / 0.29 / 0.15	6.88 / 0.19 / 0.13	6.89 / 0.20 / 0.13	6.91 / 0.20 / 0.14
	PartyScene	5.52 / 0.24 / 0.08	6.78 / 0.20 / 0.10	6.87 / 0.19 / 0.11	6.89 / 0.20 / 0.13
	BasketballDrill	6.14 / 0.22 / 0.09	6.87 / 0.20 / 0.11	6.87 / 0.20 / 0.13	6.89 / 0.20 / 0.13
	Average	5.68 / 0.28 / 0.28	6.59 / 0.20 / 0.20	6.86 / 0.20 / 0.20	6.87 / 0.20 / 0.20
Class D	*RaceHorses*	4.89 / 0.20 / 0.23	6.08 / 0.18 / 0.30	6.79 / 0.18 / 0.31	6.76 / 0.18 / 0.35
	BQSquare	7.77 / 0.16 / 0.27	6.81 / 0.18 / 0.35	6.89 / 0.17 / 0.38	6.91 / 0.17 / 0.41
	BlowingBubbles	5.16 / 0.20 / 0.23	6.79 / 0.16 / 0.29	6.87 / 0.17 / 0.37	6.89 / 0.16 / 0.40
	BasketballPass	5.63 / 0.19 / 0.23	6.30 / 0.18 / 0.28	6.83 / 0.16 / 0.32	6.83 / 0.16 / 0.35
	Average	5.68 / 0.28 / 0.28	6.59 / 0.20 / 0.20	6.86 / 0.20 / 0.20	6.87 / 0.20 / 0.20
Class F	*BaskeballDrillText*	5.12 / 0.23 / 0.08	5.77 / 0.22 / 0.09	5.77 / 0.15 / 0.10	5.79 / 0.16 / 0.13
	ChinaSpeed	5.30 / 0.28 / 0.05	6.03 / 0.16 / 0.04	6.84 / 0.18 / 0.06	6.87 / 0.15 / 0.06
	SlideEditing	5.92 / 0.22 / 0.09	6.89 / 0.15 / 0.07	6.58 / 0.15 / 0.07	6.87 / 0.16 / 0.08
	SlideShow	6.00 / 0.21 / 0.07	7.00 / 0.16 / 0.07	8.35 / 0.14 / 0.08	10.10 / 0.14 / 0.10
	Average	5.59 / 0.24 / 0.24	6.42 / 0.17 / 0.17	6.89 / 0.16 / 0.16	7.41 / 0.15 / 0.15
Average		5.71 / 0.23 / 0.14	6.50 / 0.18 / 0.16	6.86 / 0.17 / 0.18	7.04 / 0.17 / 0.20

frames in the GOP. Then, we update the D-Q and R-Q models with Eq. (10).

Step 6. If the current GOP is the last GOP in the period, go to Step 1 to start the next period (if any). Otherwise, repeat Steps 3-5 to encode the next GOP.

IV. EXPERIMENTAL RESULTS

To validate the proposed scheme, we implement the scheme on recent HM reference software, HM 16.0 [18] with RA configuration of JCTVC-L1100 [29]. The other parameters are set as default values of HM encoder. For comparison, we test four competing approaches: HEVC default RA, SVC default configuration [17] (*i.e.*, JVT-P014 [15] with $\Delta \mathrm{Qp}_{base} = 3$), Li's [22] and the proposed methods. Among them, the first two are static Qp schemes and the remaining two are adaptive schemes.

A. Comparison of RD Performance

The comparisons between the four schemes and HEVC RA environment are summarized in Table IV, where we use the HEVC RA default configuration (*i.e.*, Eq. (1) with $\Delta \mathrm{Qp}_{base} = 1$) as the benchmark and compare all the other three schemes. We measure the performance of each scheme by the average Luma Peak-Signal-to-Noise-Ratio (Y-PSNR) increase with the same Bit Rate (BR) and by the average BR increase with the same Y-PSNR. The computations are based on two criteria: Bjontegaard's average Delta PSNR (BDPSNR) and Bjontegaard's average Delta BR (BDBR) [39].

From Table IV, application of SVC QPC scheme [17] to HEVC RA (that is, increase $\Delta \mathrm{Qp}_{base}$ from 1 to 3) can barely improve the overall RD performances for most video sequences. The maximum RD improvement of this scheme is 0.0679dB in terms of BDPSNR when coding *SlideShow* of Class F and -0.83% in terms of BDBR when coding *BasketballDrill* of Class C. On the other extreme, it also causes -0.1136dB PSNR and 6.05% BDBR when coding *RaceHorses* of Class C and *BQTerrace* of Class B, respectively. On average, this scheme cause 0.0513dB BDPSNR loss and 1.95% BDBR loss.

The results indicate that: (i) the change of $\Delta \mathrm{Qp}_{base}$ in JVT-P014 [15] leads to insignificant change of coding performance; (ii) the HEVC default RA setting appears to be near-optimal for static Qp, leaving little room for further improvement without Qp adaption; (iii) an adaptive QPC scheme is needed to achieve RD performance improvements for different video sequences.

The adaptive QPC approaches, including Li's [22] (HB-4) and the proposed scheme, are able to improve the RD performances for some video sequences, as shown in Table IV. In general, Li's scheme works well for video sequences with slow motion and/or simple textures *e.g.*, *BasketballDrill* of Class C. For complex video sequences such as *RaceHorses* of Class C and *ChinaSpeed* of Class F, and high resolution video sequences in Classes A & B, it incurs large RD performance loss. The reason may be that, Li's scheme was designed for H.264 hierarchical structure, and the related model parameters were also tuned with H.264-encoded low resolution videos. However, it does fit well with HEVC encoder and high resolution video sequences. Besides, Li's scheme dynamically updates the Qps based on coding results of previous GOPs, which may introduce error propagation when an erroneous decision occurs. On average, this scheme leads to a BDPSNR of -0.0734dB and a BDBR of 5.86%.

By contrast, the proposed adaptive QPC scheme is designed for HEVC RA and results in RD improvements for most video sequences in Table IV. Although its model is derived with the one-reference frame assumption and the model parameters are tuned with Classes C & D sequences, the proposed scheme achieves robust RD improvement for all CTC classes with multiple reference frames. Even for those video sequences in which the proposed scheme achieves RD losses (such as *SteamLocomotive* of Class A and *SlideEditing* of Class F), the BDPSNR and BDBR are still almost intact as compared to the original HEVC encoder. From Table IV, the maximum BDPSNR and BDBR improvement of the proposed scheme are as high as 0.1516dB and -3.61%, respectively. On average, the proposed scheme achieves a BDPSNR of 0.0443dB and a BDBR of -1.09%. Such an improvement demonstrates the efficiency of the proposed scheme when comparing with the other QPC schemes.

To further examine the encoding efficiency of the proposed method, we compare it with the HEVC default frame-level rate control (F-RC) scheme. In general, a QPC scheme cannot achieve comparable coding performance to rate control algorithms because QPC assigns Qp from GOP level to hierarchical layer level and rate control algorithms assign Qp from video sequence level to frame and/or CU levels. However, as shown in Table IV, the proposed QPC method achieves comparable performance to the F-RC scheme and the performance can be even significantly better when coding videos of Classes A, B and F. Moreover, considering QPC aims at allocating Qs or Qps to layer level and rate control aims at allocating Qps to frame and/or CU levels, the proposed scheme may be integrated into rate control as the first step. This can be achieved by setting a threshold $\sum_{l=0}^{L-1} \Delta R_{tot,l} \leq \Delta R_{target}$ in Eq. (2) and the threshold ΔR_{target} is obtained with the target bits and predicted GOP bits without rate control. In the second step, the Qps of frames/CUs are determined through an average $\hat{Q}_l$ (as shown in Eq. (23) in each hierarchical layer.

TABLE VI

COMPUTATIONAL COMPLEXITY COMPARISONS WITH THE ORIGINAL HEVC ENCODER IN TERMS OF CODING TIME INCREASE (%)

Video		Qp=22	Qp=27	Qp=32	Qp=37
Class A	*Traffic*	0.20	-0.28	-0.28	0.29
	PeopleOnStreet	-0.50	1.18	-0.65	0.58
	Nebuta	2.02	0.56	-1.69	0.48
	SteamLocomotive	0.07	0.39	-0.28	0.50
	Average	0.45	0.46	-0.73	0.46
Class B	*Kimono*	1.72	0.72	0.38	0.47
	ParkScene	0.47	-0.22	0.60	0.01
	Cactus	-0.17	2.76	-0.33	2.18
	BQTerrace	0.57	0.43	0.07	0.22
	BasketballDrive	0.53	0.20	-0.12	0.94
	Average	0.63	0.78	0.12	0.76
Class C	*RaceHorses*	0.55	0.75	0.81	0.51
	BQMall	0.94	0.61	0.40	0.42
	PartyScene	0.72	0.58	0.19	-0.01
	BasketballDrill	0.61	0.07	-0.14	0.21
	Average	0.70	0.50	0.32	0.28
Class D	*RaceHorses*	0.96	1.33	1.14	0.64
	BQSquare	0.99	-0.01	0.38	0.46
	BlowingBubbles	0.56	0.49	-0.06	0.19
	BasketballPass	1.34	1.53	1.08	0.84
	Average	0.96	0.84	0.63	0.53
Class F	*BaskeballDrillText*	0.93	0.11	-0.03	0.03
	ChinaSpeed	0.75	1.24	1.28	0.59
	SlideEditing	0.07	0.03	-0.25	0.15
	SlideShow	0.38	0.89	0.94	0.69
	Average	0.53	0.57	0.48	0.37
Average		0.73	0.64	0.48	0.39

In addition, to intuitively show the improvement, we interpolate the RD curves first (with cubic interpolation in [39]) and calculate the Y-PSNR values of different methods at the same bit rate. The Y-PSNR improvements compared with the original encoder are presented in Fig. 2, where the first sequence of each CTC class is shown. From the figure, the proposed scheme achieves clearly better RD performances in most cases when compared with the other QPC schemes as well as the F-RC method. In conclusion, the proposed scheme outperforms the other existing QPC schemes, which demonstrates its efficiency in HEVC hierarchical coding.

B. Computational Complexity

Another significant benefit of the proposed scheme is its low complexity. We first measure the complexity of Newton-Raphson method with three terms:

TABLE VII
CONTRIBUTION OF ADAPTATION IN THE PROPOSED METHOD

Video		Gain by introducing adaptation		Fix-Rate			
		BDPSNR	BDBR	Qp = 22	Qp = 27	Qp = 32	Qp = 37
Class A	*Traffic*	0.0086	-0.23	0.56	0.00	0.00	0.00
	PeopleOnStreet	0.0310	-0.69	0.61	0.00	0.00	0.00
	Nebuta	-0.0029	-1.94	0.00	1.00	0.00	0.00
	SteamLocomotive	0.0091	-0.46	0.49	0.00	0.00	0.00
	Average	0.0115	-0.83	0.41	0.25	0.00	0.00
Class B	*Kimono*	0.0023	0.18	0.00	0.03	0.00	0.00
	ParkScene	-0.0029	0.10	0.66	0.00	0.00	0.00
	Cactus	-0.0011	0.04	0.66	0.00	0.00	0.00
	BQTerrace	-0.0014	0.05	0.61	0.01	0.00	0.00
	BasketballDrive	0.0094	-0.40	0.52	0.00	0.00	0.02
	Average	0.0013	-0.01	0.49	0.01	0.00	0.00
Class C	*RaceHorses*	0.0225	-0.60	0.68	0.19	0.00	0.00
	BQMall	0.0167	-0.41	0.47	0.01	0.00	0.00
	PartyScene	0.0262	-0.62	0.63	0.03	0.00	0.00
	BasketballDrill	0.0571	-1.33	0.02	0.00	0.00	0.00
	Average	0.0306	-0.74	0.45	0.06	0.00	0.00
Class D	*RaceHorses*	0.0331	-0.70	0.65	0.00	0.00	0.03
	BQSquare	0.0085	-0.24	0.69	0.01	0.00	0.00
	BlowingBubbles	0.0141	-0.33	0.58	0.02	0.00	0.00
	BasketballPass	0.0292	-0.59	0.21	0.00	0.00	0.05
	Average	0.0212	-0.46	0.53	0.01	0.00	0.02
Class F	*BaskeballDrillText*	0.0441	-0.95	0.03	0.00	0.00	0.00
	ChinaSpeed	0.0402	-0.75	0.44	0.02	0.02	0.00
	SlideEditing	-0.0223	0.14	0.00	0.14	0.24	0.00
	SlideShow	-0.0043	0.03	0.29	0.34	0.61	0.55
	Average	0.0144	-0.38	0.19	0.12	0.22	0.14
Average		0.0151	-0.46	0.42	0.09	0.04	0.03

AveNumIterPerGOP, represents the average number of iterations when perform QPC in one GOP; AveTimePerIter, represents the average time (in terms of μs) per iteration; and TotIterTimePPM, represents the total iteration time in terms of part per million of the overall encoding time. The test was run on a 2011 Laptop with a Quad-Core i7-2.70GHz CPU and 8GB memory. The measure results are summarized in Table V. It can be seen that on average, less than 10 times of Newton-Raphson iterations are required to perform QPC and the average time of each iteration is less than 1 μs. As the Qp increase, the AveNumIterPerGOP slightly increases. Besides, the total time cost of these iterations is less than 1 part per million of the overall encoding time. Therefore, the computation overhead of Newton-Raphson method is negligible.

The computational complexity of the overall proposed scheme is also studied. It can be generally changed due to both the computational overhead and changes of coding Qps. In Table VI, we use the coding time increase to measure the complexity as

$$T_{inc} = \frac{T_p - T_o}{T_o} \times 100\%, \tag{27}$$

where T_p and T_o denote the encoding time of the proposed scheme and the original HEVC encoder, respectively. To reduce randomness in the measurement, we run all experiments three times and average the coding time of all tests to estimate T_p and T_o.

From Table VI, the computational complexity has barely been increased after introducing the proposed adaptive QPC scheme. It can also be observed that for some video sequences, the proposed approach runs faster than the original encoder. This is because the computational overhead of the proposed approach is negligible. Investigations of the Newton-Raphson method also show that, less than 10 iterations are required for most cases, which ensures a low computational complexity. Therefore, the coding time differences between the original HEVC and the proposed approaches are mostly determined by the changing of Qps, which may be positive or negative. Therefore, our scheme may even reduce the overall coding time in some cases.

C. Contribution of Adaptation

In order to measure the contribution of the adaptation in the proposed scheme, we compare the coding performances of the proposed adaptive QPC scheme and its degenerate version without adaptation. The RD gains with adaptation are summarized in Table VII. It can be seen that, the proposed scheme is able to greatly improve the RD performance for a range of video sequences when introducing the Qp adaptation. This robust adaptation feature once again indicates that an adaptive QPC method is preferable. On average, by introducing the adaptation, additional 0.0151 dB BDPSNR and -0.46% BDBR can be achieved. The adaptation also makes the proposed scheme more practical for real-world QPC adaptation for diverse types of video sequences.

We also check the probability when the adaptive QPC scheme cannot work well for which static QPC is turned on, as we discussed in Step 3 of Section III.B. The probabilities of all the tested video sequences are presented in Table VII, where we can see that the probability is usually very low except when a low Qp is employed. There may be two reasons why this may happen. First, the initial values in $\hat{\mathbf{x}}$ are selected empirically and the parameters of RD models are coarsely predicted from

TABLE VIII
COMPARISON OF DIFFERENT METHODS WITH HEVC LD

Video		SVC [17]		Li's [22]		F-RC		Proposed-Static	
		BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR
Class C	*RaceHorses*	-0.1673	4.41	-0.2738	7.34	0.1072	-2.71	-0.0769	2.01
	BQMall	-0.0142	0.38	-0.0413	0.98	0.1059	-2.71	0.0268	-0.69
	PartyScene	0.1202	-2.79	0.1584	-3.66	0.1435	-3.52	0.1073	-2.53
	BasketballDrill	0.0890	-2.22	0.1858	-4.33	0.1841	-4.68	0.1200	-2.94
	Average	0.0069	-0.05	0.0073	0.08	0.1352	-3.41	0.0443	-1.03
Class D	*RaceHorses*	-0.1141	2.40	-0.2459	5.92	-0.0010	0.01	-0.0423	0.90
	BQSquare	0.0683	-1.87	0.1176	-3.31	-0.0288	0.83	0.0725	-2.03
	BlowingBubbles	0.0794	-1.93	0.0721	-1.70	0.1458	-3.56	0.0746	-1.82
	BasketballPass	-0.0854	1.71	-0.0945	1.99	0.2590	-5.18	0.0032	-0.07
	Average	-0.0130	0.08	-0.0377	0.72	0.0937	-1.97	0.0270	-0.76
Class F	*BaskeballDrillText*	0.0956	-2.21	0.2212	-4.82	0.2114	-5.00	0.1305	-2.99
	ChinaSpeed	-0.0988	1.94	0.1096	-2.03	0.2939	-6.05	0.1159	-2.27
	SlideEditing	-0.3191	2.11	-0.2224	1.53	-4.7459	43.51	0.1892	-1.28
	SlideShow	0.2703	-3.42	0.4420	-5.45	-0.6894	10.99	0.1257	-1.58
	Average	-0.0130	-0.40	0.1376	-2.69	-1.2325	10.86	0.1403	-2.03
Average		-0.0109	0.09	0.0701	-1.97	-0.3009	2.67	0.0803	-1.85

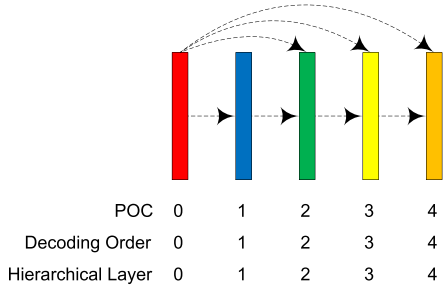


Fig. 3. An example of an LD prediction structure with a GOP size of 4.

the previous GOP, which may lead to zero-valued determinant or divergence in the Newton-Raphson method. Second, when the predicted layer quantization values $\hat{Q}_l, l = 0, 1, \ldots, L-1$ are very close to each other, the rounding process may result in identical Qp value in all layers. To retain the inter-layer differences and ensure the RD performance, we choose to use the static QPC scheme for both cases, as we have discussed in Sections III.B.

D. Testing on LD and LP Coding

Technically the proposed scheme can also be applied in HEVC LD/LP encoders, however with a lower RD gain than RA. By directly testing the proposed scheme with video sequences of Classes C, D and F, the average BDPSNR and BDBR of LD and LP are 0.0248/-0.48 and 0.0233/-0.43, respectively. This is not surprising because LD/LP prediction structures are very different from that in RA. An example of LD structure is presented in Fig. 3. Compared with Fig. 1, we can see that in LD structure, there are 2 reference pictures in one direction and only 1 of them is from the neighboring layer of the current frame/slice. Therefore in LD/LP configurations, the proposed scheme does not have the space to operate on as compared to the RA case.

Instead of exploiting the complex inter-layer dependencies of LD/LP, we design Qp assignment methods between LD/LP GOPs, as discussed in Section II.C. This generally works because there are multiple successive GOP structures in an intra period of LD/LP. Following the idea of Section II.C, we still utilize Eq. (13), where L denotes the number of successive GOPs, $\Delta D_{tot,l}$ and $\Delta R_{tot,l}$ represent the distortion and bit rate increments of a video sequence when changing the Qp of GOP l, respectively. Experiments also show linear D dependency and zero R dependency between LD/LP GOPs. Therefore, it can be easily obtained for inter-GOP Qp assignments:

$$\Delta D_{tot,l} = \frac{1-\delta_{GOP}^{L-l}}{1-\delta_{GOP}} \cdot \Delta D_l, \tag{28}$$

$$\Delta R_{tot,l} \approx \Delta R_l, \tag{29}$$

where $\delta_{GOP} = 0.48$ denotes the error propagation between GOPs; ΔD_l and ΔR_l represent the distortion and bit rate increments of GOP l, respectively.

The above inter-GOP Qp assignment problem can be solved with the procedures given in Section III. A notable difference is that, inter-GOP QPC is conducted before encoding any GOP, and thus it is actually not an "adaptive" scheme. The comparisons between the proposed static scheme and other schemes are presented in Tables VIII and IX for LD and LP coding, respectively. From the tables, the proposed static scheme is superior to SVC scheme and F-RC; it is also almost comparable to Li's adaptive scheme, especially for LD coding. In the future, we will try to explore the intra-GOP LD/LP dependencies to derive adaptive QPC schemes for LD/LP.

E. Additional Analysis on QPC

We study how the quality fluctuation changes after adopting the proposed adaptive QPC scheme. We define a measure Σ, which indicates the standard variation of all Qps within a GOP. A smaller Σ indicates a smoother video quality. In the proposed scheme, $\Sigma < 3$ because we have clipped the Qps of all layers into $[\bar{Q}_p - 3, \bar{Q}_p + 3]$, as discussed in Section III.A.

Three tests are then performed:

1) Test how Σ adapts to different $\bar{Q}$s. Fig. 4 shows the impact of $\bar{Q}$ on the average Σ of all GOPs, where 4 Class C sequences and 4 Class D sequences are presented. In general, Σ decreases when $\bar{Q}$ increases, *i.e.*, the optimal Qp set with

TABLE IX
COMPARISON OF DIFFERENT METHODS WITH HEVC LP

Video		SVC [17]		Li's [22]		F-RC		Proposed-Static	
		BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR	BDPSNR	BDBR
Class C	*RaceHorses*	-0.1603	4.27	-0.1738	4.78	0.1382	-3.57	-0.0718	1.89
	BQMall	-0.0174	0.49	0.0009	-0.09	0.1131	-3.00	0.0240	-0.64
	PartyScene	0.1026	-2.54	0.1289	-3.19	0.1346	-3.51	0.0936	-2.35
	BasketballDrill	0.0741	-1.89	0.1762	-4.25	0.2038	-5.27	0.1124	-2.80
	Average	-0.0002	0.08	0.0330	-0.69	0.1474	-3.84	0.0395	-0.97
Class D	*RaceHorses*	-0.1232	2.65	-0.1554	3.82	0.0137	-0.30	-0.0438	0.96
	BQSquare	0.0744	-2.32	0.1124	-3.86	0.0142	-0.42	0.0791	-2.50
	BlowingBubbles	0.0762	-1.92	0.0759	-1.91	0.1345	-3.45	0.0726	-1.86
	BasketballPass	-0.0805	1.61	-0.0166	0.39	0.2616	-5.25	0.0084	-0.18
	Average	-0.0133	0.01	0.0041	-0.39	0.1060	-2.35	0.0291	-0.89
Class F	*BaskeballDrillText*	0.0832	-1.99	0.2005	-4.63	0.2300	-5.61	0.1254	-2.97
	ChinaSpeed	-0.0914	1.79	0.1661	-3.20	0.2999	-6.14	0.1199	-2.36
	SlideEditing	-0.4113	2.85	-0.1458	0.99	-4.6552	41.09	0.0380	-0.28
	SlideShow	0.2831	-3.62	0.5274	-6.63	-0.6315	10.29	0.1209	-1.55
	Average	-0.0341	-0.24	0.1871	-3.37	-1.1892	9.91	0.1010	-1.79
Average		-0.0182	0.12	0.1008	-2.68	-0.2749	1.90	0.0685	-1.84

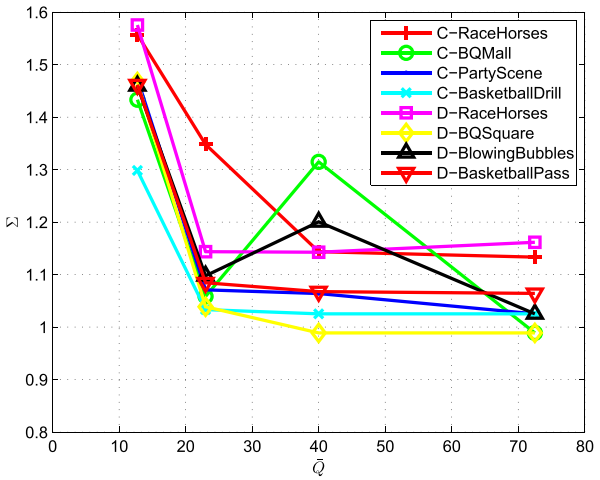
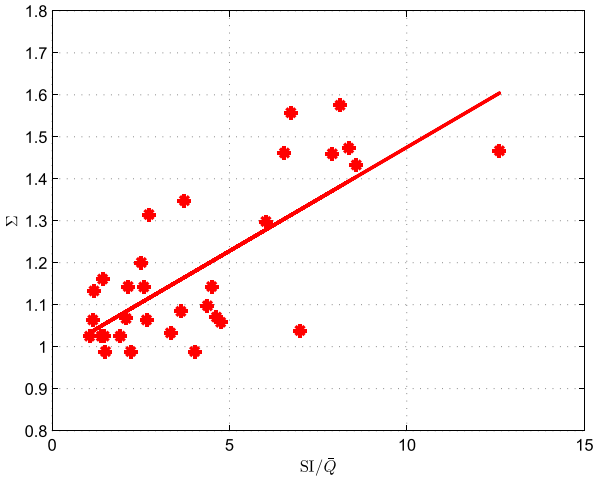


Fig. 4. The proposed QPC method adapts to different $\bar{Q}$ values.

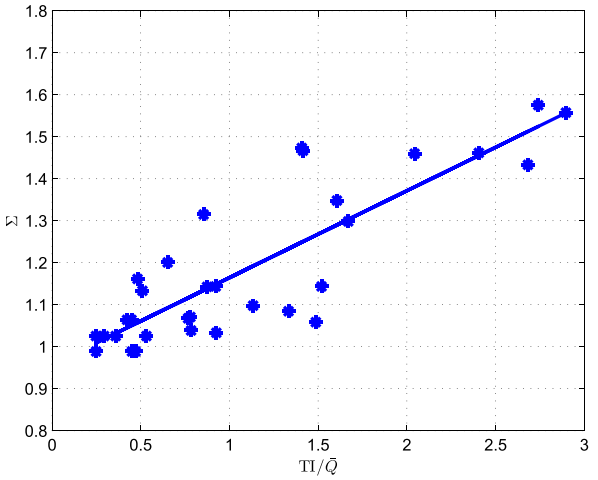
a smaller $\bar{Q}$ has more diverse Qps than that with a larger $\bar{Q}$. This fact can also provide guidance on how to further improve the static QPC scheme. For example, in the QPC scheme of Eq. (1), $\Delta \mathrm{Qp}_{base}$ should be larger for smaller Qps.

2) Test how Σ adapts video contents with different textures and motions. Here we use Spatial Information (SI) and Temporal Information (TI) [40] to measure the complexity of textures and motions, respectively. The relationships between the average Σ, $\mathrm{SI}/\bar{Q}$ and $\mathrm{TI}/\bar{Q}$ are illustrated in Fig. 5, using the data from Fig. 4. We notice that Σ can be approximated as an linearly increasing function of either $\mathrm{SI}/\bar{Q}$ or $\mathrm{TI}/\bar{Q}$. In other words, with the same $\bar{Q}$, a video with complex texture and motion prefers an optimal Qp set with diverse Qps. This fact provides insights on how to further improve both the adaptive and static QPC scheme in the future.

3) Test how Σ adapts in real-time. We present the update processes of four benchmark video sequences in Fig. 6, where the Σ of GOP 0 is determined by the initial Qp set and the Σ values of the following GOPs are determined by the proposed adaptive QPC scheme. In this figure, the period of Σ shows the intra period of video sequence. For example, the



(a)



(b)

Fig. 5. The proposed QPC method adapts to different SI and TI values. (a) Σ versus $\mathrm{SI}/\bar{Q}$. (b) Σ versus $\mathrm{TI}/\bar{Q}$.

intra period of *RaceHorses* of Class D is 32, therefore, the period of Σ is 4 GOPs. It is observed that in each RA period, Σ could quickly converge to different values in the first two

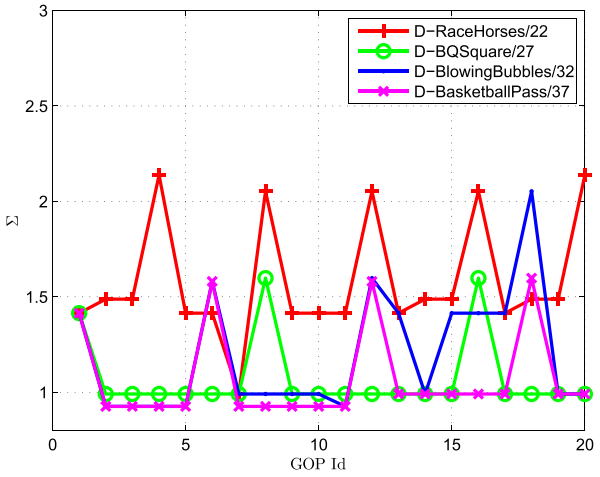


Fig. 6. The proposed QPC method adapts in timeline.

or three GOPs. That is, the adaptive QPC scheme provides a smoother video quality with less quality fluctuation.

V. Conclusions and Future Work

We present an adaptive QPC scheme to improve the HEVC hierarchical video coding with RA. To the best of our knowledge, this is the first effort contributed to developing an adaptive QPC scheme for HEVC. Simulation results have demonstrated the high efficiency, low complexity and fast adaptation of the proposed scheme. In particular, we have formulated the QPC problem as a non-linear programming problem that also applies to other Qp assignment problems in HEVC. The additional analysis presented in Section IV.E shall also benefit both static and adaptive QPC approaches for other types of applications.

In the proposed model, we use δ to illustrate the inter-hierarchical layer dependencies and we adopt a constant δ for all video sequences and encoder settings for simplicity. Additional improvement is expected if an estimated or dynamically updated δ is adopted for a given sequence with different encoder settings. How to determine δ in a single-pass algorithm remains as an open issue. Another possible improvement lies in a better understanding of inter-layer RD dependencies of LD/LP encoders, as discussed in Section V.D. We will study these issues in our future work.

References

[1] M. T. Pourazad, C. Doutre, M. Azimi, and P. Nasiopoulos, "HEVC: The new gold standard for video compression: How does HEVC compare with H.264/AVC?" *IEEE Consum. Electron. Mag.*, vol. 1, no. 3, pp. 36–46, Jul. 2012.

[2] 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.

[3] B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand, *High Efficiency Video Coding (HEVC) Text Specification Draft 10 (for FDIS & Consent)*, document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Jan. 2013.

[4] B. Li, J. Xu, D. Zhang, and H. Li, "QP refinement according to Lagrange multiplier for high efficiency video coding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2013, pp. 477–480.

[5] M. Wang, K. N. Ngan, H. Li, and H. Zeng, "Improved block level adaptive quantization for high efficiency video coding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 509–512.

[6] L. Zhao, X. Zhang, Y. Tian, R. Wang, and T. Huang, "A background proportion adaptive Lagrange multiplier selection method for surveillance video on HEVC," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2013, pp. 1–6.

[7] S. Li, C. Zhu, Y. Gao, Y. Zhou, F. Dufaux, and M.-T. Sun, "Lagrangian multiplier adaptation for rate-distortion optimization with inter-frame dependency," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 117–129, Jan. 2016.

[8] H. Zeng, K. N. Ngan, and M. Wang, "Perceptual adaptive lagrangian multiplier for high efficiency video coding," in *Proc. IEEE Picture Coding Symp. (PCS)*, Dec. 2013, pp. 69–72.

[9] S. Wang, S. Ma, D. Zhao, and W. Gao, "Lagrange multiplier based perceptual optimization for high efficiency video coding," in *Proc. Annu. Summit Conf. Asia-Pacific Signal Inf. Process. Assoc. (APSIPA)*, Dec. 2014, pp. 1–4.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[11] J. Qi, X. Li, F. Su, Q. Tu, and A. Men, "Efficient rate-distortion optimization for HEVC using SSIM and motion homogeneity," in *Proc. Picture Coding Symp. (PCS)*, Dec. 2013, pp. 217–220.

[12] F. Cen, Q. Lu, and W. Xu, "SSIM based rate-distortion optimization for intra-only coding in HEVC," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2014, pp. 17–18.

[13] S. Ma, S. Wang, S. Wang, L. Zhao, Q. Yu, and W. Gao, "Low complexity rate distortion optimization for HEVC," in *Proc. Data Compress. Conf. (DCC)*, Mar. 2013, pp. 73–82.

[14] S. Guo, Z. Liu, D. Wang, Q. Han, and Y. Song, "Linear rate estimation model for HEVC RDO using binary classification based regression," in *Proc. Data Compress. Conf. (DCC)*, Mar. 2014, p. 406.

[15] H. Schwarz, D. Marpe, and T. Wiegand, *Hierarchical B Pictures*, document JVT-P014, Poznań, Poland, Jul. 2005.

[16] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical B pictures and MCTF," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2006, pp. 1929–1932.

[17] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[18] K. McCann, C. Rosewarne, B. Bross, M. Naccari, K. Sharman, and G. J. Sullivan, *High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description*, document JCTVC-R1002, ITU-T/ISO/IEC, Jun./Jul. 2014.

[19] L. Jing, L. Ma, and C. Zhang, "Improved strategy of QP decision for hierarchical B pictures structure in SVC," *Proc. SPIE*, vol. 8761, pp. 1–6, Mar. 2013.

[20] W. Wan, Y. Chen, Y.-K. Wang, M. M. Hannuksela, H. Li, and M. Gabbouj, "Efficient hierarchical inter picture coding for H.264/AVC baseline profile," in *Proc. Picture Coding Symp. (PCS)*, May 2009, pp. 1–4.

[21] X. Li, P. Amon, A. Hutter, and A. Kaup, "Model based analysis for quantization parameter cascading in hierarchical video coding," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 3765–3768.

[22] X. Li, P. Amon, A. Hutter, and A. Kaup, "Adaptive quantization parameter cascading for hierarchical video coding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May/Jun. 2010, pp. 4197–4200.

[23] H. L. Cycon *et al.*, "alisch, "Adaptive temporal scalability of H.264-compliant video conferencing in heterogeneous mobile environments," in *Proc. Global Telcomm. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–5.

[24] H. L. Cycon, T. C. Schmidt, M. Wälisch, D. Marpe, and M. Winken, "A temporally scalable video codec and its applications to a video conferencing system with dynamic network adaption for mobiles," *IEEE Trans. Consum. Electron.*, vol. 57, no. 3, pp. 1408–1415, Aug. 2011.

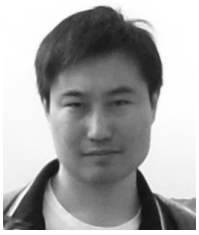
[25] J. Liu, Y. Cho, and Z. Guo, "Single pass dependent bit allocation for H.264 temporal scalability," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep./Oct. 2012, pp. 705–708.

[26] J. Liu, Y. Cho, and Z. Guo, "Single-pass dependent bit allocation in temporal scalability video coding," in *Proc. Data Compress. Conf. (DCC)*, Mar. 2013, p. 507.

[27] S. Sanz-Rodríguez and T. Schierl, "A rate control algorithm for HEVC with hierarchical GOP structures," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2013, pp. 1719–1723.

[28] T. Zhao, Z. Wang, and S. Kwong, "Flexible mode selection and complexity allocation in high efficiency video coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1135–1144, Dec. 2013.

[29] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-L1100, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Geneva, Switzerland, Jan. 2013.
[30] S. Hu, H. Wang, S. Kwong, T. Zhao, and C.-C. J. Kuo, "Rate control optimization for temporal-layer scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1152–1162, Aug. 2011.
[31] H. Wang and S. Kwong, "Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 140–144, Jan. 2008.
[32] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.
[33] N. Kamaci, Y. Altunbasak, and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 994–1006, Aug. 2005.
[34] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 12, pp. 1533–1544, Dec. 2005.
[35] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
[36] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 186–199, Feb. 1999.
[37] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1983.
[38] Y. Liu, Z. G. Li, and Y. C. Soh, "A novel rate control scheme for low delay video communication of H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 1, pp. 68–78, Jan. 2007.
[39] G. Bjontegaard, *Calculation of Average PSNR Differences Between RD-Curves*, document VCEG-M33, Austin, TX, USA, Apr. 2001.
[40] *Subjective Video Quality Assessment Methods for Multimedia Applications*, document ITU-T Rec. P.910, International Telecommunication Union, 1999.



Tiesong Zhao (S'08–M'12) received the B.S. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2006, and the Ph.D. degree in computer science from the City University of Hong Kong, Hong Kong, in 2011. He is currently a Qishan Scholar with the College of Physics and Information Engineering, Fuzhou University, Fuzhou, China. From 2011 to 2012, he was a Research Associate with the Department of Computer Science, City University of Hong Kong. Until 2013, he served as a Post-Doctoral Research Fellow with the Department Electrical and Computer Engineering, University of Waterloo, ON, Canada. From 2014 to 2015, he was a Research Scientist with the Department of Computer Science and Engineering, State University of New York at Buffalo, NY, USA. His research interests include image/video processing, visual quality assessment, video coding and transmission.

Zhou Wang (S'99–M'02–SM'12–F'14) received the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, in 2001. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include image processing, coding, and quality assessment; computational vision and pattern analysis; multimedia communications; and biomedical signal processing. He has more than 100 publications in these fields with over 24 000 citations (Google Scholar). He was a member of the IEEE Multimedia Signal Processing Technical Committee (2013-2015). He served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING (2009-2014), *Pattern Recognition* (2006- present), and the IEEE SIGNAL PROCESSING LETTERS (2006-2010), and a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2013-2014 and 2007-2009), the *EURASIP Journal of Image and Video Processing* (2009-2010), and *Signal, Image and Video Processing* (2011-2013). He was a recipient of the 2014 NSERC E.W.R. Steacie Memorial Fellowship Award, the 2013 IEEE Signal Processing Best Magazine Paper Award, the 2009 IEEE Signal Processing Society Best Paper Award, the 2009 Ontario Early Researcher Award, and the ICIP 2008 IBM Best Student Paper Award (as senior author).

Chang Wen Chen (F '04) received the B.S. degree from the University of Science and Technology of China in 1983, the M.S.E.E. degree from the University of Southern California in 1986, and the Ph.D. degree from the University of Illinois at Urbana–Champaign in 1992. He is currently an Empire Innovation Professor of Computer Science and Engineering with the University at Buffalo, State University of New York. He was an Allen Henry Endow Chair Professor with the Florida Institute of Technology from 2003 to 2007. He was a Faculty Member of Electrical and Computer Engineering with the University of Rochester from 1992 to 1996 and the University of Missouri–Columbia from 1996 to 2003.

He has been the Editor-in-Chief of the IEEE TRANSACTIONS ON MULTIMEDIA since 2014. He served as the Editor-in-Chief of the EEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2006 to 2009. He has been an Editor for several other major IEEE TRANSACTIONS and JOURNALS, including the PROCEEDINGS OF IEEE, the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, and the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS. He has served as the Conference Chair for several major IEEE, ACM, and SPIE conferences related to multimedia, video communications, and signal processing. His research is supported by NSF, DARPA, Air Force, NASA, Whitaker Foundation, Microsoft, Intel, Kodak, Huawei, and Technicolor.

He and his students have received eight Best Paper Awards or Best Student Paper Awards over the past two decades. He has also received several research and professional achievement awards, including the Sigma Xi Excellence in Graduate Research Mentoring Award in 2003, the Alexander von Humboldt Research Award in 2010, and the State University of New York at Buffalo Exceptional Scholar—Sustained Achievement Award in 2012. He is an SPIE Fellow.