This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

Joint Machine Learning and Game Theory for Rate Control in High Efficiency Video Coding

Wei Gao, *Member*, *IEEE*, Sam Kwong, *Fellow*, *IEEE*, and Yuheng Jia

***Abstract*—In this paper, a joint machine learning and game theory modeling (MLGT) framework is proposed for inter frame coding tree unit (CTU) level bit allocation and rate control (RC) optimization in High Efficiency Video Coding (HEVC). First, a support vector machine (SVM) based multi-classification scheme is proposed to improve the prediction accuracy of CTU-level Rate-Distortion (R-D) model. The legacy "chicken-and-egg" dilemma in video coding is proposed to be overcome by the learning-based R-D model. Second, a mixed R-D model based cooperative bargaining game theory is proposed for bit allocation optimization, where the convexity of the mixed R-D model based utility function is proved, and Nash bargaining solution (NBS) is achieved by the proposed iterative solution search method. The minimum utility is adjusted by the reference coding distortion and frame-level Quantization parameter (QP) change. Lastly, intra frame QP and inter frame adaptive bit ratios are adjusted to make inter frames have more bit resources to maintain smooth quality and bit consumption in the bargaining game optimization. Experimental results demonstrate that the proposed MLGT based RC method can achieve much better R-D performances, quality smoothness, bit rate accuracy, buffer control results and subjective visual quality than the other state-of-the-art one-pass RC methods, and the achieved R-D performances are very close to the performance limits from the FixedQP method.**

***Index Terms*—H.265/HEVC; video coding; rate control; bit allocation; machine learning; support vector machine (SVM); R-D model classification; game theory; Nash bargaining solution (NBS); non-numerical solution; iterative search.**

I. Introduction

HIGH Efficiency Video Coding (H.265/HEVC) [1] has become the most popular video coding standard after its final release in 2013 [2] by the Joint Collaborative Team on Video Coding (JCTVC), which is organized by the ITU-T Video Coding Experts Group (VCEG) and the ISO Moving Picture Experts Group (MPEG). HEVC achieves better coding efficiency than its predecessor H.264/AVC [3] by adopting many new and complex coding tools [4]. As high performance video coding can effectively relieve the burdens of massive video storage and transmission, it acts as a key role in practical video streaming and multimedia communication systems.

A. Rate Control Problem in Video Coding

As one of the key techniques in video coding and transmission, rate control (RC) endeavors to provide the best coding performance by controlling the quantization process. The optimization techniques for RC have been widely investigated [5]-[17], which encodes group-of-pictures (GOP) [5], frames [6] and coding tree units (CTU) [7], [31] by properly setting Quantization parameters (QPs) under a given bit rate constraint with the goal of achieving better RC performances. In general, the optimization objectives of RC mainly include better R-D performances, accurate bit rate achievements, visual quality smoothness and stable buffer control to avoid overflow and underflow cases [6].

The existing works on RC optimization are mainly focused on the following aspects [5]-[17]: (1) for different coding structures, such as all intra (AI) [7], low delay B and P (LB/LP) [8], and random access (RA) [6]; (2) for different optimization objectives, such as R-D performance [6]-[7], quality smoothness and buffer control [6], [8]-[10]; (3) at different coding levels, such as GOP-level [5], frame-level [6], and block-level [7]. There are also some RC works on the perceptual quality optimization for video coding [40] by exploiting efficient phase information [41]. The analysis strategies and optimization methods for RC can be divided into three categories: (1) *Q*-domain based RC methods [11]-[12], which achieve better R-D performances by exploring the relationship between Quantization step (Q_{step}) and D, as well as the relationship between Q_{step} and R. For HEVC, the related typical work is the pixel-wise Unified Rate Quantization (URQ) model based RC [11]-[12]; (2) λ-domain based RC methods [6], [13]-[16], which jointly consider the Lagrange multiplier decisions for both Rate-Disortion-Optimization (RDO) and RC. For HEVC, the related typical work is the R-λ model based RC [13]-[15], which uses the fixed QP-λ relationship [16] and has been adopted by JCTVC and implemented in the latest HM16.8 [17]; (3) ρ-domain based RC methods [5]-[6], [10], which are based on the assumption that the percentage of non-zero quantized Discrete Cosine Transform (DCT) coefficients is highly correlated with R, D, and Q_{step}, respectively.

Moreover, there are also some efforts on using two-pass RC for video coding performance improvements [38]-[39]. However, the main drawback is the doubled computation overhead, which may prevent it from real-time video

This work was supported in part by Natural Science Foundation of China under Grants 61672443, in part by Hong Kong RGC General Research Fund 9042489 (CityU 11206317) and in part by Hong Kong RGC General Research Fund 9042322 (CityU 11200116).

Wei Gao is with the National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University, Shenzhen 518060, China, and is also with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China. (e-mail: weigao5-c@my.cityu.edu.hk).

Sam Kwong and Yuheng Jia are with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, and Sam Kwong is also with City University of Hong Kong Shenzhen Research Institute, Shenzhen 5180057, China (e-mail: cssamk@cityu.edu.hk; yhjia3-c@my.cityu. edu.hk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

transmission. As the well-known "chicken-and-egg" dilemma in video coding, the coding complexity of prediction residuals at CTU-level is greatly influenced by QP. Obviously, different QPs will generate different mode decisions and motion estimation strategies, where we will have different coding complexities for prediction residuals at CTU-level, as well as the corresponding R-D relationships. Then the collected data and parameters in the first pass are not always accurate or useful for reference in the second pass, especially for CTUs with different video content. Therefore, many existing works [6]-[16], [40] make efforts to overcome the "chicken-and-egg" dilemma in one-pass RC algorithms in order to achieve more accurate R-D modelings and better coding performances.

In [18], the traditional parameter updating method in R-λ model cannot effectively handle the cases with drastic motions or abrupt scene changes, e.g. screen content videos. In fact, even for camera captured videos, the drastic motions or abrupt scene changes are also very common at CTU-level. However, in the parameter updating for R-λ model [13], the video content change is assumed to be very smooth. Therefore, the traditional parameter updating methods, which are based on the assumption of smooth content change, cannot effectively improve the CTU-level bit allocation and RC performances.

To overcome the legacy "chicken-and-egg" dilemma in video coding to achieve accurate R-D models, almost all the existing RC methods adopt the spatial-temporal prediction methods for the updating of R-D model (or the equivalent R-λ model) parameters, which are not always accurate, especially for CTU coding blocks. But the achievement of accurate R-D models is the prerequisite to achieve the improvements on R-D performance by RC. Additionally, as a key step of RC, there is no doubt that bit allocation needs an effective modeling to achieve optimization goals. Therefore, an accurate prediction method for CTU-level R-D model and an effective modeling strategy for bit allocation optimization are both required in a desirable RC scheme.

B. Machine Learning for Rate Control Optimization

Fortunately, in the current big data era, machine learning is an effective tool to provide an accurate model. There are already some works using machine learning to solve fast mode decision (MD) problems in video coding to obtain complexity reductions [19]. However, very few works can be found on using learning for RC optimization. In [20], frame-level QP increments are predicted by a radial basis function (RBF) network for quality smoothness of H.264/SVC in variable bit rate (VBR) environment. The extracted RC related features include buffer fullness, consumed bits, buffer size, and target buffer fullness. The limitation of this work is that it is proposed for quality smoothness with VBR, not for R-D performance improvements and constant bit rate (CBR). In our work, R-D performance optimization and quality smoothness will be jointly considered in CTU-level RC optimization with CBR.

Bit allocation in RC greatly influences R-D performances and quality smoothness, and most of bit allocation methods are based on R-D models [6]-[7]. As we know, because of the drastic motions and abrupt scene changes, CTU-level coding complexity changes are very large even for adjacent frames, therefore, the traditional prediction and updating methods based on spatial and temporal information for CTU-level R-D model parameters will not be accurate.

In this paper, in order to achieve accurate R-D models, a machine learning based R-D model prediction method is proposed for CTU-level bit allocation and RC optimization. A novel multi-classification scheme based on support vector machine (SVM) is proposed to predict the CTU-level R-D models. The features are extracted from the original video data and the coding results of previous frames. The proposed machine learning based method can achieve more accurate CTU-level R-D models than the traditional prediction methods, especially for CTUs with drastic motions and abrupt scene changes. In addition, we will also implement a scene change detection based R-D model prediction method, and the coding results will be compared to further verify the effectiveness of the learning-based approach.

C. Game Theory for Rate Control Optimization

Game theory is an effective modeling approach to solve resource allocation problems, while bit allocation in RC is exactly a resource allocation problem. Therefore, there are a few of works using game theory for bit allocation optimization [7], [21]-[22]. Generally, game theory has two different categories, i.e. non-cooperative games [33] and cooperative games [34]. In [21], Nash equilibrium [35] can be achieved in the non-cooperative game for the bit allocation modeling among multiple objects in a frame. In [7] and [22], cooperative games are modeled for bit allocation optimization among different CTUs in HEVC and MPEG-4, respectively. Cooperative games [34] are very suitable for bit allocation modeling, in which each coding frame or unit can be deemed as a player, and then Nash bargaining solution (NBS) [23] can be achieved. Different CTUs and frames can be modeled as players to compete for bit resources for their own coding utility improvements [7]. However, in the existing cooperative game modelings [7], [22], only single R-D model based bit allocation optimization problem is discussed, and the mixed R-D model cases are not mentioned. Moreover, the minimum utility is defined only from the coding distortion or consumed bits of the previous collocated frames or CTUs [7], which is not enough to take into account different optimization goals together, e.g. better R-D performances and smoother quality.

In this paper, we will provide a discussion on the convexity of the R-utility mapping function based on mixed R-D models in the feasible utility set, and propose an iterative solution search method for the non-numerical solution problem in the CTU-level bargaining game where mixed R-D models are applied. Moreover, the minimum utility in the bargaining game is also refined by the joint consideration on the reference coding distortion and the frame-level QP. This will help maintain good coding quality smoothness and make the CTU-level bit allocation competition much more reasonable. The minimum utility is also incorporated into the analysis to achieve the cubic function solution.

D. Contributions and Organization of This Paper

As shown in Fig. 1, a joint machine learning and game theory modeling (MLGT) based RC optimization framework is proposed to improve the inter frame CTU-level RC. The proposed MLGT based RC mainly consists of two steps. In the first step, accurate CTU-level R-D models can be achieved by using machine learning. In the second step, bit allocation optimization and QP determination can be achieved by the proposed cooperative bargaining game theory modeling. The learning based R-D model prediction is implemented by a multiclass SVM classifier after effective feature extraction, and then the NBS solution is derived for the cooperative game theory based bit allocation optimization by using the proposed iterative solution search method. After the CTU-level QP is determined, the encoding process is conducted, including MD, motion estimation (ME), motion compensation (MC), DCT transform and quantization, and entropy coding which is implemented by context-based adaptive binary arithmetic coding (CABAC) to generate compressed bit streams.

The main contributions of the proposed MLGT based RC optimization framework can be listed as follows:

(1) Adopting machine learning technique to overcome the legacy "chicken-and-egg" dilemma in video coding by providing accurate R-D model prediction: We investigate the R-D relationships for inter frame CTUs and present the model parameter ranges. Then, a multi-classification scheme based on SVM is proposed to improve the prediction accuracy for CTU-level R-D models, which can get rid of the traditional spatial-temporal based prediction methods. Moreover, to verify the effectiveness of the proposed learning-based method, we also apply the scene change detection based RC parameter resetting method for CTU-level R-D model prediction and compare the coding performances.

(2) Mixed R-D model based cooperative game theory modeling for CTU-level resource allocation problem: We propose a mixed R-D model based cooperative game theory method to optimize the bit allocation performances. For the mixed R-D models, the convexity of R-utility mapping function in the feasible utility set is proved for the proposed bargaining game. To overcome the problem that closed-form solutions cannot be derived for the bargaining game, an iterative solution search method is proposed. The minimum utility definition is also refined by the joint consideration on the reference coding distortion and frame-level QP, which helps maintain smooth coding quality and make CTU-level bit competition much more reasonable.

Moreover, intra frame QP determination and adaptive bit ratios among frames are also proposed to be refined. More bits are allocated to inter frames to best exploit the potentials of the proposed MLGT RC method. Much closer frame-level bit ratios can be helpful for coding quality smoothness.

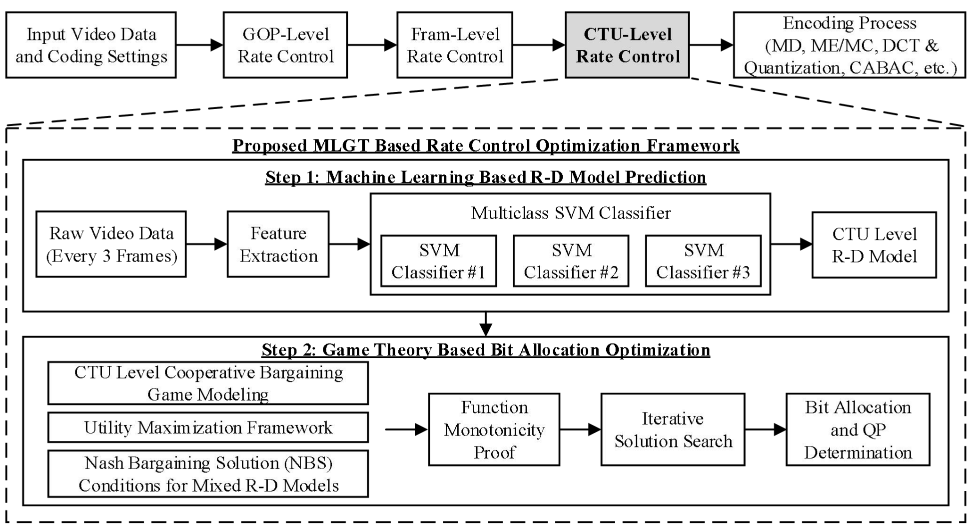


Fig. 1. The proposed joint machine learning and game theory (MLGT) based rate control optimization framework.

The remainder of this paper is organized as follows. In Section II, the R-D relationship for inter frame CTUs is investigated, and the SVM-based multi-classification scheme is proposed for accurate R-D model prediction. The scene change detection based CTU-level R-D model prediction method is proposed to highlight the necessity of using machine learning. In Section III, cooperative game theory based bit allocation optimization scheme is proposed for inter frame CTUs, where the utility function is defined and the convexity of mapping function in the feasible utility set is proved. In Section IV, the NBS solution is provided and an iterative solution search method is proposed for the mixed R-D model based bargaining game. The final CTU-level bit allocation, QP determination and the overall RC algorithm are also given in this section. Experimental results are presented in Section V, and conclusions are drawn in Section VI.

II. OVERCOME "EGG-AND-CHICKEN" DILEMMA BY MACHINE LEARNING BASED R-D MODEL PREDICTION

A. R-D Modeling for Inter Frame CTUs

The CTU-level RC scheme from JCTVC-M0036 [15] has been implemented in HM16.8 [17], where the allocated bits for CTUs are based on the bit weights derived from frame-level λ and related R-λ model parameters. The model parameters will be updated by the actual consumed bits R and actual λ values of the collocated CTUs in the previous frame. Nevertheless, the problem is that the block-level coding complexity can easily vary even for adjacent frames. The drastic motions and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

abrupt scene changes in CTUs will greatly reduce the robustness and accuracy of the adopted R-λ model. Therefore, the traditional spatial-temporal prediction method cannot work well for CTU-level R-D modelings. In addition, the collocated CTUs in adjacent several frames may use same or close QPs, which easily makes the real-time regression fail.

Obviously, the accuracy improvement on CTU-level R-D modeling will benefit bit allocation and RC performances a lot. To achieve accurate CTU-level R-D models with large coding complexity variations, we propose to use machine learning to predict the R-D relationship based on the information from the previous and current changes of original video signals. First of all, effective features are suggested to be extracted from the previous three frames, and then a machine learning algorithm can be devised to identify which kind of R-D model can be used for the current CTU with the best modeling accuracy.

It should be noted that there are a lot of SKIP modes in inter frames [6], which have unique R-D relationships. SKIP mode avoids the coding process and directly copies the coded CUs as the prediction for the current CU. The bit consumption of SKIP mode only comes from the flag information while the distortion is highly related to the referenced pixels. Therefore, skipped CUs have a quite different R-D relationship from the other non-skipped CUs. In this paper, CTUs with a lot of skip modes are denoted as "SKIP_Most_CTUs" or "SMC", and this kind of CTUs should be specially handled. In the experiments, the CTU-level consumed bits per pixel (bpp) is used to determine SKIP_Most_CTU. The CTUs with bit consumption less than the threshold of 20 bits (i.e. 0.005 bpp) are deemed as SKIP_Most_CTUs.

According to the Cauchy distribution model for transformed DCT coefficients [24], the R-D model exhibits the following power function based relationship

$$R_i = m_i D_i^{-n_i}, \tag{1}$$

where m_i and n_i are model parameters, and m_i, n_i >0. Therefore, besides SKIP_Most_CTUs, the other CTUs can be analyzed by using the power function based R-D model. To test the CTU-level R-D relationship in inter frames, an experiment has been conducted, where five QPs (QP=20, 26, 32, 38, 44) are used for quantization process.

In Table I, it can be seen that the fitting accuracy can be up to 96.14% in average, and the average values of n_i are in the range of (0.5871, 1.9696). Since a classification approach is proposed in this paper to discriminate different CTUs, all the other CTUs except for SKIP_Most_CTUs are proposed to be classified into two categories for simplicity. The CTUs with the following 1.0-order and 1.5-order R-D models are deemed as two categories

$$R_i = C_{1,i} D_i^{-1.0}, \tag{2}$$

$$R_i = C_{2,i} D_i^{-1.5}, \tag{3}$$

where R_i and D_i are the consumed bits and mean-squared-error (MSE) based coding distortion, respectively. $C_{1,i}$ and $C_{2,i}$ are the model parameters which can reflect the coding complexity of the predicted residual signals.

In fact, it is inadvisable to divide CTUs into too many classes. The reason is that the computational complexity of multiclass classifier will be very high which is not suitable for real-time video coding, and also a high classification accuracy of multiclass classifier cannot be guaranteed for many classes. Therefore, in this paper all CTUs are proposed to be classified into three categories, which is enough for the accurate R-D modelings for CTUs and has a relatively low complexity for the real-time classification task.

In Fig. 2, for PeopleOnStreet, the CTUs with index CTUID=16 and CTUID=23 prefer to select the 1.0-order and 1.5-order R-D models, respectively, while for BasketballDrill, the CTUs with index CTUID=21 and CTUID =54 prefer to select the 1.0-order and 1.5-order R-D models, respectively. It can be seen that different CTUs will select different R-D models for better curve fitting results. This validates the necessity of using different R-D models for different CTUs.

TABLE I
PARAMETER RANGES AND FITTING ACCURACY FOR POWER FUNCTION BASED R-D MODELS

Class	Sequence	m_{min}	m_{max}	n_{min}	n_{max}	Accuracy
A	PeopleOnStreet	0.1925	13.5626	0.5234	2.1960	0.9876
	Traffic	0.0697	38.5621	0.6006	1.6329	0.9668
B	BasketballDrive	0.1771	136.8781	0.6163	3.7048	0.9394
	BQTerrace	0.0148	49.2892	0.1528	1.9362	0.9473
	Cactus	2.8172	82.4321	0.7385	2.7570	0.9642
	Kimono	0.1410	19.0636	0.5725	2.6896	0.9704
	ParkScene	0.0437	25.1527	0.6581	2.9199	0.9670
C	BasketballDrill	0.2555	15.2360	0.6782	1.6391	0.9836
	BQMall	0.3750	16.1885	0.6219	1.9062	0.9781
	PartyScene	0.5907	54.9573	0.5215	1.5404	0.9609
	RaceHorsesC	0.0930	44.1109	0.5939	1.6799	0.9713
D	BasketballPass	0.3135	18.4179	0.5723	1.2517	0.9808
	BlowingBubbles	2.0962	29.2897	0.7092	1.6099	0.9759
	BQSquare	0.9374	20.4423	0.5817	1.3138	0.9614
	RaceHorses	0.3160	30.8347	0.5699	1.6758	0.9815
E	FourPeople	0.0542	5.2208	0.6978	1.4086	0.9596
	Johnny	0.0707	2.7832	0.5445	1.8092	0.8829
	KristenAndSara	0.0851	7.8171	0.6151	1.7814	0.9263
	Average	**0.4802**	**33.9021**	**0.5871**	**1.9696**	**0.9614**

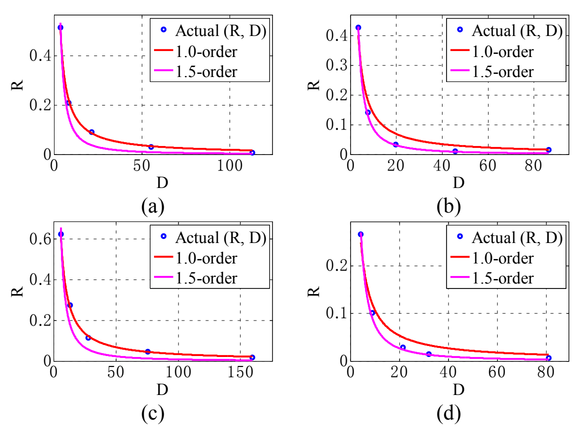


Fig. 2. CTU-level R-D relationship modeling. (a) PeopleOnStreet, CTUID=16; (b) PeopleOnStreet, CTUID=23; (c) BasketballDrill, CTUID=21; (d) BasketballDrill, CTUID=54.

The 18 sequences from Class A to Class E have been tested, and the R-D relationships of inter frame CTUs can be observed. The curve fitting results for 1.0-order and 1.5-order models are compared in Table II. The two models can perform differently for different sequences and for different CTUs. The 1.0-order model can perform much better than the 1.5-order model for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

some sequences and CTUs, and vice verse. If the 1.0-order and 1.5-order power function based R-D models are separately used for all CTUs, the fitting accuracy results are 95.17% and 92.23%, respectively. These two kinds of power orders both have the possibility to perform better than the other. If the best model is selected for each CTU, the selected percentages for these two models are 21.32% and 7.47%, respectively, and the other 71.20% CTUs select SKIP_Most_CTUs. The total curve fitting accuracy can be increased to 97.27%. Therefore, it will be helpful to design a classification scheme to select the best R-D model for each CTU to improve the fitting accuracy.

TABLE II
SEPARATE R-D MODELS AND OPTIMAL SELECTION PERCENTAGES

Class	Fitting Accuracy		Selected Percentages			Best Accuracy
	1.0-order	1.5-order	1.0-order	1.5-order	SMC	
A	0.9738	0.9425	0.2955	0.1023	0.6022	0.9869
B	0.9225	0.9330	0.0690	0.1295	0.8015	0.9646
C	0.9655	0.9140	0.3217	0.0947	0.5837	0.9743
D	0.9573	0.8547	0.3679	0.0161	0.6161	0.9580
E	0.9395	0.9676	0.0122	0.0311	0.9567	0.9796
Average	**0.9517**	**0.9223**	**0.2132**	**0.0747**	**0.7120**	**0.9727**

It should be noted that the fitting accuracy improvement is very important for each CTU individual. For the R-D model based RC optimization, it is well recognized that accurate R-D modelings for CTUs can increase the RC performances. But the inaccurate R-D modelings and bit allocations for a few of CTUs will degrade the total coding performances significantly. The reason is that the remaining bits and allocated bits may be sensitively influenced by other CTUs. Moreover, we should treat SKIP_Most_CTUs in a special way, which takes a large majority of all CTUs. Therefore, we have reasons to believe that the increased accuracy for R-D modelings will benefit the further RC optimization. To improve the curve fitting accuracy, it is necessary to devise a multiclass classifier to discriminate different CTUs with different R-D relationships, including SKIP_Most_CTUs, 1.0-order and 1.5-order R-D models.

B. Scene Change Based R-D Model Prediction and Necessity of Using Machine Learning Scheme

Since there are always a lot of drastic motions and abrupt scene changes at CTU-level, it is difficult to achieve accurate R-D models by using the real-time spatial-temporal prediction updating method. It is needed to propose an more efficient method for accurate R-D model prediction.

It is well known that the detection for scene changes can be used for intra frame insertion and adaptive GOP length decision [36]-[37], which will reset RC parameters by using intra coding. These inserted intra frames can avoid serious coding quality degradation and fluctuations due to abrupt scene changes in a video sequence. For the CTU-level scene change detection, the frame-level correlation measurement using difference of histogram (DOH) in [36] is proposed to be adopted for CTUs in this paper. The CTU-level DOH can be calculated as the absolute sum of the histogram difference between the CTU block B_i in the i-th frame and its collocated CTU block B_{i+1} in the $(i+1)$-th frame

$$DOH\left(B_i, B_{i+1}\right)=\sum_{j=0}^{L}\left|h_i\left(j\right)-h_{i+1}\left(j\right)\right|, \tag{4}$$

where h_i and h_{i+1} are histograms of the CTU blocks B_i and B_{i+1}, respectively, and L is the maximum gray level. For 8 bit depth video data, L is equal to 255. For fairness, the histograms are normalized by the number of the CTU block pixels.

To investigate the relationship between the scene change metric DOH and CTU-level R-D model selection, we have conducted an experiment to collect the coding results of CTUs. For the 18 video sequences, the CTUs in five frames of each sequence are encoded. The total number of tested CTUs is 28990. In Fig. 3, different R-D model selections are marked, including 1.0-order, 1.5-order and SKIP_Most_CTU. It can be seen that there are a lot of overlapped DOH ranges, which indicates that the correlation between DOH and R-D model selection is very low. We can use the thresholds for scene change measurements to decide the R-D model category. Obviously, the determination accuracy is easily influenced by the threshold settings.

We use the brute force approach to find the best thresholds to distinguish different R-D models on the total 28990 CTUs. The achieved optimal DOH thresholds, DOH_1 and DOH_2, are 0.3783 and 0.2409, respectively. The CTUs with DOHs above DOH_1 are determined as 1.5-order R-D model, while those with DOHs below DOH_2 are determined as SKIP_Most_CTUs. The others are determined as 1.0-order R-D model. The overall R-D modeling accuracy is 0.7500. However, it should be noted that, for inter frames the SKIP_Most_CTUs take a large majority of the total CTUs. For the result of 0.7500, {0.0637, 0.0141, 0.6722} are from the 1.0-order, 1.5-order and SKIP_Most_CTU R-D models, respectively. As the optimal selected percentages of different R-D models shown in Table II, the classification accuracy for each class is very low. Because each individual of CTU-level bit allocation greatly influences others, more accurate R-D modelings for all CTUs are expected to improve bit allocation and RC optimization.

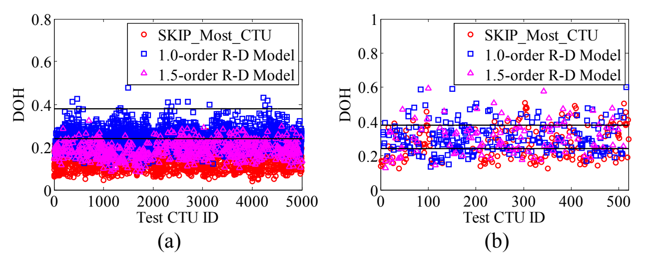


Fig. 3. R-D model prediction using the derived thresholds based on scene change measurement DOH. (a) PeopleOnStreet; (b) RaceHorsesC.

In addition, in Fig. 3 (a) and (b), when the achieved optimal thresholds are applied to PeopleOnStreet and RaceHorsesC, it can be seen that they cannot effectively distinguish different R-D model selections for CTUs. Therefore, based on the above results and analysis, the R-D model change cannot be simply achieved by the detection results for scene changes.

To better handle this problem, it is much more advisable to use machine learning to achieve more accurate R-D models. Moreover, the extracted feature DOH for scene change is not enough effective, while more useful features need to be extracted for better prediction accuracy.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

The above analysis provides a scene change based R-D model prediction method using DOH feature, which will be jointly implemented with the proposed game theory based bit allocation scheme. The scene change based method is denoted as "Proposed SCGT", and its performances will be tested and compared with the proposed MLGT method.

C. Feature Extraction for R-D Model Classification

For R-D model prediction problem, features that reflect the coding complexity should be considered. To extract efficient features from the original YUV signals, the difference maps are calculated for every adjacent three frames.

We firstly get the absolute difference map (ADM) ΔF_i from the current frame F_i and its previous frame F_{i-1}, where i indicates the frame coding order, and then the ADM ΔF_{i-1} from the previous frame F_{i-1} and the second previous frame F_{i-2}. The selected features include the mean and variance values for CTUs in the ADMs ΔF_i and ΔF_{i-1}, and in the gradient maps of the ADMs (GADMs). For each CTU, features will be extracted from the two ADMs and two GADMs.

The following features are extracted, which are highly related to the coding complexity in R-D model: (1) the mean of ADMs; (2) the standard variance of ADMs; (3) the square root of the mean of the squared ADMs; (4) the square root of the standard variance of the squared ADMs; (5) the mean of GADMs; (6) the standard variance of GADMs; (7) the square root of the mean of the squared GADMs; (8) the square root of the standard variance of the squared GADMs. For each YUV component, these 8 features are extracted from the two ADMs. A flag is also used to indicate whether the consumed bits of the previous collocated CTU is below a threshold, which is a feature highly related to SKIP_Most_CTUs. Therefore, in total 49-dimensional features are used for each CTU.

Moreover, we can give a conceptual analysis for the reason that features should be extracted from two difference maps. We denote the CTU-level difference between current frame and its previous frame as $\Delta diff_1$ for the first previous frame, and $\Delta diff_2$ for the second previous frame. If $\Delta diff_1$ is large, it is not sure that the coding complexity is large, since there is a possibility that the motion estimation can work well to remove coding redundancy and then the predicted residuals have low coding complexity. Thus, a single $\Delta diff_1$ can not represent the coding complexity well. If $\Delta diff_1$ and $\Delta diff_2$ are both large, motion estimation may work well, and the coding complexity can be low. If $\Delta diff_1$ is large and $\Delta diff_2$ is small, the coding complexity may be large. If $\Delta diff_1$ is small and $\Delta diff_2$ is large, the coding complexity may be small. If $\Delta diff_1$ and $\Delta diff_2$ are both small, the coding complexity may be small. Therefore, two difference maps from three frames have a better representational ability and can be helpful for better learning of the R-D model selection.

D. Multiclass SVM Classifier

As an effective tool for binary classification problem [25], the SVM classifier attempts to obtain the margin maximization for the classification hyperplane. The optimal hyperplane $f(x)=\omega^T\phi(x)+b=0$ is defined to be well trained to best discriminate samples, where x is input sample vector, ϕ is a kernel function, which maps x into a higher dimensional space and introduces the non-linear transformation to improve the learning performance. The bias term is denoted as b, which can adjust the flexibility of the hyperplane.

In υ-SVM based binary classification [26], the optimal hyperplane is achieved by minimizing the cost function

$$\begin{gathered} J_1(\omega,b,\rho,\xi)=\frac{1}{2}\|\omega\|^2-\upsilon\psi+\frac{1}{L}\sum_{i=1}^{L}\varepsilon_i \\ s.t.\ \ y_i\left(\omega^T\varphi(x_i)+b\right)\geq\psi-\varepsilon_i,\ \varepsilon_i\geq 0,\ i=1,2,...,L,\ \psi\geq 0 \end{gathered} \quad (5)$$

where x_i is the i-th samples, y_i={-1, +1} is the target label, ε_i is the non-zero soft margin and slack variable. The margin is adjusted by the parameter ψ, which makes the two classes have the margin $2\psi/\|\omega\|$. Parameter υ can influence the prediction accuracy and the number of selected support vectors, and is set to 0.05 in the experiments. The Gaussian radial basis function (RBF) based kernel function is used.

The final determined label is achieved by checking the sign of decision function value (DFV), which evaluates the input sample based on the obtained support vectors and bias term b. One of the advantages for using υ-SVM classifier is that the number of support vectors can be adjusted by parameter υ.

Although the binary SVM classifier only can handle the two-class classification problem, it can also be extended to handle the multiclass cases for the R-D model prediction problem in this paper. DFV can be used to measure the distance between input sample and hyperplane. Therefore, the DFV distance based one-versus-rest scheme can be adopted for the multi-classification problem [27]. In this paper, three binary SVM classifiers are used to discriminate which R-D model the current CTU belongs to. Before that, we should devise three SVM classifiers #1, #2 and #3 to predict whether the current CTU belongs to 1.0-order R-D model, 1.5-order R-D model and SKIP_Most_CTUs, respectively. By collecting the DFV value in each classification, the final decision can be made for the R-D model selection.

The used three classifiers #1, #2 and #3 can generate three labels $Label_1$, $Label_2$ and $Label_3$, and three decision function values DFV_1, DFV_2 and DFV_3, respectively. From [27], the final multiclass decision label FDL can be achieved as

$$FDL=\arg\max_{i=1,2,3} DFV_i-1, \quad (6)$$

which means that the classifier with the largest DFV can give the final decision label in the multi-classification problem.

E. Classification Accuracy

The inter frame CTUs from 9 video sequences, including PeopleOnStreet, BasketballDrive, Cactus, ParkScene, PartyScene, RaceHorsesC, BlowingBubbles, RaceHorses, Johnny, are encoded for training data collection, while the CTUs from the other 9 video sequences are used for testing.

In Table III, the classification accuracy can be up to 87.46%, 88.84% and 94.34% for Classifier #1, #2 and #3, respectively. When the three classifiers are combined to fulfill the multi-classification task based on (6), the overall accuracy can be up to 86.07%. Classifier #1, #2 and #3 are binary classifiers for a separate R-D model type classification, while the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

multiclass classifier gives a multi-classification result. The adopted multiclass SVM classifier can ensure that most of CTUs are correctly predicted to have accurate R-D models, which is expected to benefit further RC performances.

TABLE III
CLASSIFICATION ACCURACY OF THREE SEPARATE BINARY CLASSIFIERS AND MULTICLASS SVM CLASSIFIER

Class	Classifier #1	Classifier #2	Classifier #3	Multiclass Classifier
A	0.8090	0.9570	0.9310	0.8450
B	0.8481	0.9392	0.8873	0.8373
C	0.9039	0.7885	0.9279	0.8173
D	0.8393	0.7679	1.0000	0.8393
E	0.9729	0.9896	0.9708	0.9646
Average	**0.8746**	**0.8884**	**0.9434**	**0.8607**

In the proposed joint MLGT RC optimization framework, the multiclass SVM classifier is proposed for accurate CTU-level R-D model prediction, while the game theoretical modeling approach is proposed for bit allocation optimization. In the following sections, we will give detailed discussions on the cooperative bargaining game for CTUs using the proposed mixed R-D models and its NBS solution will be derived for bit allocation optimization.

III. COOPERATIVE GAME MODELING FOR BIT ALLOCATION

A. Bargaining Game for Inter Frame CTUs

CTU-level bit allocation in inter frames can be modeled as a cooperative game, where each CTU can be deemed as a player in the bargaining game. Every CTU player competes for more bit resources for encoding to improve its own coding quality. The constraint for the bit competition is the target bits for the current frame. Excessive bit consumptions of some CTUs will greatly impede the bit achievements of other CTUs, which will influence the overall coding quality and utility in the current frame. Therefore, every CTU player should be rational in the bargaining game, and all CTU players are advised to work cooperatively to compete for bit resources to obtain the overall coding quality and utility gains in the current frame.

For video coding, the utility definition should be related to the coding distortion or quality, e.g. the structural similarity metric (SSIM) as in [7]. Therefore, to maximize the overall utility, the R-D relationships for CTUs need to be considered comprehensively. First of all, the proposed learning-based scheme can improve the CTU-level R-D modeling accuracy. Then, due to the different R-D characteristics owned by different CTUs, different bit increments may be required for the same distortion reduction. Therefore, for better utility achievement, a cooperative bargaining game is proposed in this paper to model the CTU-level bit allocation for inter frame, where different R-D characteristics are taken into account.

In the bargaining game, each CTU player i (i=1, 2, ... , N) competes for bit resources rationally to improve its own utility $U_i(r_i)$, where r_i is the consumed bits of the i-th CTU, and the bit competition can further improve the overall utility of the current frame. The ultimate bargaining result is that the overall utility reaches its optimization limit. Different CTU players have different bargaining powers due to the different R-D models and coding complexities. In the bargaining process, each player would not yield without a limit. The minimum utilities $\{U_{i,d}(r_{i,d}), i=1, 2, \dots, N\}$ should be guaranteed in the bargaining game for each player, where $r_{i,d}$ is the lowest bit consumption to achieve the minimum utility $U_{i,d}$. Afterwards, the remaining bits are used for the further bargaining. These minimum utilities are also named disagreement points, which may be quite different for different CTU players.

The utilities $\{U_i(r_i), i=1, 2, \dots, N\}$ for all CTUs in a frame can define a feasible utility set U, while the minimum utilities $\{U_{i,d}(r_{i,d}), i=1, 2, \dots, N\}$ can defined the minimum utility set U_d. All CTU players with $\{U_i(r_i)\}$ and $\{U_{i,d}(r_{i,d})\}$ form a N-player cooperative bargaining game, in which the Pareto optimal solution $U^*=\{U^*_i(r_i), i=1, 2, \dots, N\}$ can be obtained by NBS [7], [23], and the overall utility can be maximized. Here, the NBS solution needs to be in the feasible utility set. For each CTU player, it plays rationally to ensure the minimum utility, which guarantees $U^*_i > U_{i,d}$. Therefore, in the bargaining game, the achieved NBS based bit allocation scheme can provide the Pareto optimality.

The existing works [7], [22] both adopt single R-D model based bargaining game for bit allocation optimization. As we know, the accurate R-D modeling is the prerequisite condition to improve RC performances. However, the single R-D model based bit allocation is not always enough accurate, therefore the mixed R-D models are adopted for bit allocation optimization in this paper. To the best of our knowledge, there are no works using mixed R-D model based game theory modeling approach for bit allocation and RC optimization.

B. Bargaining Utility Definition

Except for SKIP_Most_CTUs, two categories of R-D models have been used for other CTUs, including 1.0-order and 1.5-order R-D models as shown in (2) and (3). The utility definition for the i-th CTU can be expressed as

$$U_i = 1/D_i, \tag{7}$$

where D_i is MSE-based coding distortion for the i-th CTU. The utility mapping function f: $U_i \to r_i$ can be built to depict the relationship between the utility U_i and the consumed bits r_i. From (2) and (3), all utility mapping functions belong to the following two types

$$r_i = f_1(U_i) = C_{1,i} U_i, \tag{8}$$

$$r_i = f_2(U_i) = C_{2,i} U_i^{3/2}, \tag{9}$$

where each CTU selects one of the mapping functions.

In the bargaining game, every CTU will not compromise to reduce its own utility without a limit, and owns a disagreement point which is the minimal guarantee on the achieved utility. The corresponding utility for the disagreement point of the i-th CTU can be expressed as

$$U_{i,d} = 1/D_{i,d}, \tag{10}$$

where $D_{i,d}$ is the maximum allowed coding distortion for the i-th CTU in the bargaining game. From (8) and (9), the consumed bits $r_{i,d}$ to achieve the utility can be expressed as

$$r_{i,d} = f_1(U_{i,d}) = C_{1,i} U_{i,d}, \tag{11}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

$$r_{i,d} = f_2(U_{i,d}) = C_{2,i} U_{i,d}^{3/2}. \tag{12}$$

respectively.

C. Adjusted Minimum Utility Determination

In the proposed bargaining game, the minimum utility $U_{i,d}$ can be calculated as

$$U_{i,d} = \gamma U_{i,pre,adjust} = \gamma / (\delta D_{i,pre}), \tag{13}$$

where γ is set to a constant 0.5, the utility $U_{i,pre}$ of the i-th CTU in the previous frame is adjusted to be $U_{i,pre,adjust}$, and the corresponding distortion for $U_{i,pre}$ is denoted as $D_{i,pre}$. The adjusting factor δ can be roughly defined as

$$\delta = Q_{step,cur} / Q_{step,pre}, \tag{14}$$

where $Q_{step,cur}$ is the Q_{step} of current frame, and $Q_{step,pre}$ is the Q_{step} of previous frame with the same temporal level (TL). Thus, a more reasonable minimum utility definition method is proposed by the joint consideration on the reference coding distortion and frame-level QP change. This will help maintain coding quality smoothness and make the CTU-level bit allocation competition much more reasonable.

D. Mixed R-D Model Based Utility Function Convexity

Before adopting NBS to solve the bargaining game, in the feasible utility set U, the mixed R-D model based utility mapping function should satisfy the convexity condition [28].

Theorem 1**:** The utility mapping function f: $U_i \rightarrow r_i$ (as shown in (8) and (9)) is convex in the feasible utility set U for the mixed R-D model based cooperative bargaining game.

This theorem can be proved as **Appendix A**. It can seen that the convexity of the utility function can be satisfied separately for each CTU. Furthermore, because the utility mapping functions f_1 and f_2 in (8) and (9) are both monotonically increasing functions, all the possible achieved utilities can be larger than the minimum utility at the disagreement point. After the convexity of mapping function is proved, the NBS can be applied for bit allocation optimization.

IV. Mixed R-D Model Based Nash Bargaining Solution

A. D-Q and R-Q Models

In [6], it has been proved that the distortion D has a linear relationship with Q_{step} in HEVC encoding

$$Q_{step} = k_1 D + k_2, \tag{15}$$

where k_1 and k_2 are model parameters, k_2 is very close to zero. Let k=k_1, then we can have

$$Q_{step} = kD. \tag{16}$$

Take the R-D models (2) and (3) into account, then we have the following R-Q_{step} relationships

$$R_i = k_i C_{1,i} / Q_{step,i}, \tag{17}$$

$$R_i = k_i^{3/2} C_{2,i} / Q_{step,i}^{3/2}, \tag{18}$$

for the two different R-D models, respectively.

B. Mixed R-D Model Based Nash Bargaining Solution

By adopting the proposed SVM multi-classification scheme, all inter frame CTUs are classified into three categories, including SKIP_Most_CTUs, 1.0-order and 1.5-order R-D models. As aforementioned, SKIP_Most_CTUs have special R-D characteristics, for which the target bits are proposed to be predicted from the collocated CTUs in the previous frame with the same TL. As SKIP_Most_CTUs are excluded from the bargaining game optimization, the remaining bits R_c is calculated as

$$R_c = R_f - \sum_{i=1}^{N_{SMC}} r_{SMC,i}, \tag{19}$$

where R_f is the target bits for the current frame, $R_{SMC,i}$ is the predicted bits for the i-th SKIP_Most_CTU, N_{SMC} is the number of SKIP_Most_CTUs, R_c will be allocated to CTUs with 1.0-order and 1.5-order R-D models.

When encoding the current frame, we have a R-D model order map *RDOrder_Map* from the SVM multi-classification scheme, and also have C_1 and C_2 parameter map *C_Map* from the previous frame. These two maps are highly related to the CTU-level coding complexity.

Every CTU has its own utility and the total utility of current frame can be maximized in the bargaining process

$$\max \prod_{i=1}^{N} (U_i - U_{i,d}),\ s.t. \begin{cases} r_i \ge r_{i,d} \\ \sum_{i=1}^{N} r_i \le R_c \\ r_{i,\min} \le r_i \le r_{i,\max} \end{cases}, \tag{20}$$

where the frame-level utility can be maximized, r_i and $r_{i,d}$ are the total allocated bits and minimum utility bits for the i-th CTU, respectively, while U_i and $U_{i,d}$ are the achieved utility and the minimum utility, respectively. R_c is the total bit constraint, and $r_{i,\min}$ and $r_{i,\max}$ are the minimum and maximum allowable bits for the i-th CTU, respectively.

The multiplication maximization can be transformed to the summation maximization in the logarithm domain

$$\max \sum_{i=1}^{N} \ln(U_i - U_{i,d}),\ s.t. \begin{cases} r_i \ge r_{i,d} \\ \sum_{i=1}^{N} r_i \le R_c \\ r_{i,\min} \le r_i \le r_{i,\max} \end{cases}. \tag{21}$$

Solve it by introducing Lagrange multipliers

$$\begin{aligned} J = &\sum_{i=1}^{N} \ln(U_i - U_{i,d}) + \xi\left(\sum_{i=1}^{N} r_i - R_c\right) + \sum_{i=1}^{N} \varphi_i (r_i - r_{i,d}) \\ &+ \sum_{i=1}^{N} \theta_{1,i}(r_i - r_{i,\min}) + \sum_{i=1}^{N} \theta_{2,i}(r_{i,\max} - r_i) \end{aligned}, \tag{22}$$

where J can reach to its maximum if the following Karush-Kuhn-Tucker (KKT) conditions [29] are solved

$$\begin{cases} \dfrac{\partial J}{\partial r_i} = \dfrac{\partial \ln(U_i - U_{i,d})}{\partial r_i} + \xi + \varphi_i + \theta_{1,i} + \theta_{2,i} = 0 \\ \dfrac{\partial J}{\partial \varphi_i} = r_i - r_{i,d} \geq 0 \\ \varphi_i \dfrac{\partial J}{\partial \varphi_i} = \varphi_i (r_i - r_{i,d}) = 0 \\ \dfrac{\partial J}{\partial \xi} = \sum_{i=1}^{N} r_i - R_c \leq 0 \\ \dfrac{\partial J}{\partial \theta_{1,i}} = \dfrac{\partial}{\partial \theta_{1,i}} \left(\theta_{1,i} (r_i - r_{i,\min}) \right) = 0 \\ \dfrac{\partial J}{\partial \theta_{2,i}} = \dfrac{\partial}{\partial \theta_{2,i}} \left(\theta_{2,i} (r_{i,\max} - r_i) \right) = 0 \end{cases} . \tag{23}$$

Since the consumed bits r_i is in the range of ($r_{i,\min}$, $r_{i,\max}$), $\theta_{1,i}=\theta_{2,i}=0$. Similarly, because $r_i>r_{i,d}$, φ_i should also be 0. Then, from the first equation in (23), we can have

$$\frac{\partial \ln(U_i - U_{i,d})}{\partial r_i} + \xi = 0 \cdot \tag{24}$$

Based on the U_i and $U_{i,d}$ definitions in (7) and (10), we can see that only U_i is correlated with r_i, while $U_{i,d}$ is independent of r_i. Since the achieved utility for each CTU player must be larger than the minimum utility at the disagreement point in the bargaining game, U_i-$U_{i,d}$>0. Then, from (24)

$$\frac{\partial \ln(U_i - U_{i,d})}{\partial r_i} + \xi = \frac{1}{U_i - U_{i,d}} \cdot \frac{\partial U_i}{\partial r_i} + \xi = 0 \cdot \tag{25}$$

In (8) and (9), two types of mapping functions can describe the relationship between U_i and r_i. If the current CTU selects the 1.0-order or 1.5-order R-D models, r_i can be calculated as

$$r_i = C_{1,i} U_{i,d} - 1/\xi , \tag{26}$$

$$r_i - C_{2,i}^{2/3} U_{i,d} r_i^{1/3} + 2/(3\xi) = 0 , \tag{27}$$

respectively. In (27), the cubic function can be solved by adopting the Cardano's method [30]. In **Appendix B**, the solution for r_i in (27) is given and the monotonic relationship between r_i and ξ is analyzed based on video coding practices and the proposed minimum disagreement settings.

For simplicity, the relationships between ξ and r_i in (26) and (27) are denoted as $r_i=h_1(\xi)$ and $r_i=h_2(\xi)$, respectively. Whether a CTU selects 1.0-order or 1.5-order R-D model, the summation of consumed bits in a frame should satisfy the bit constraint requirement in (20). The numbers of CTUs that select the 1.0-order and 1.5-order R-D models are denoted as N_1 and N_2, respectively, then $N_1+N_2=N$. If CTUs are grouped based on the two categories, we can have

$$R_{cal} = \sum_{m=1}^{N_1} r_{1,m} + \sum_{n=1}^{N_2} r_{2,n} = \sum_{m=1}^{N_1} h_{1,m}(\xi) + \sum_{n=1}^{N_2} h_{2,n}(\xi) \leq R_c , \tag{28}$$

where R_{cal} is the bits calculated from the Lagrange multiplier ξ, and is expected to approximate R_c by adaptively setting ξ.

To get the optimal bit allocation scheme for CTUs according to (26) or (27), the Lagrange multiplier ξ should be firstly achieved to meet the bit consumption constraint. However, this is a non-numerical solution problem. In this paper, a new iterative solution search method is proposed to obtain ξ for all CTUs. Afterwards, the optimal CTU-level bit allocation scheme can be given.

C. Proposed Iterative Solution Search Method

From (2) and (3), $C_{1,i}>0$ and $C_{2,i}>0$. Therefore, r_i and ξ in (26) exhibit a monotonically increasing relationship. From the bargaining game rule, $U_i>U_{i,d}$, then

$$1/\xi = C_{1,i} U_{i,d} - r_i < C_{1,i} U_i - r_i = 0 , \tag{29}$$

therefore $\xi<0$.

In **Appendix B**, it can be proved that, for (27), the range of ξ is $-3^{1/2}/r_{i,d}<\xi<0$, and r_i and ξ also exhibit a monotonically increasing relationship in the range of $r_i>r_{i,d}$.

Therefore, R_{cal} and ξ exhibit a monotonically increasing relationship in (28). The monotonic relationship is the prerequisite condition for the following proposed iterative solution search method, which finds the best ξ to make the total consumed bits R_{cal} approximate the target bits R_c for the CTUs with 1.0-order and 1.5-order R-D models in a frame. The best ξ is searched in the estimated range of $-3^{1/2}/r_{i,d}<\xi<0$. As the previous frame already has the values of R_i, and $C_{1,i}$ or $C_{2,i}$, we can get a set of ξ_i values. The minimum and maximum ξ values are denoted as $\xi_{\min}$ and $\xi_{\max}$, respectively. They are also scaled as

$$\xi_{\min} \leftarrow \xi_{\min} / s_{\min} , \tag{30}$$

$$\xi_{\max} \leftarrow s_{\max} \xi_{\max} , \tag{31}$$

where $s_{\min}$ and $s_{\max}$ are scaling factors and both set as 10 in the experiment. Afterwards, we use the proposed iterative search algorithm to find the best ξ for the frame-level bit constraint.

Based on the ξ value, we can get the calculated consumed bits from the following function $g(\xi)$

$$g(\xi) = \sum_{m=1}^{N_1} h_{1,m}(\xi) + \sum_{n=1}^{N_2} h_{2,n}(\xi) \cdot \tag{32}$$

We can have $g(\xi_{\min})=R_{\min}$ and $g(\xi_{\max})=R_{\max}$, respectively. The mean of $\xi_{\min}$ and $\xi_{\max}$ is denoted as ξ_{mid}. The half division operation for the search range is conducted in the iteration search process. The iteration time is recorded as Iter and its maximum value is denoted as Iter_max. After each iteration search, the search range [$\xi_{\min}$, $\xi_{\max}$] becomes smaller. In the experiments, Iter_max is set to 20, then the precision of the achieved ξ can reach to $(\xi_{\max}-\xi_{\min})/2^{20}$. The monotonically increasing characteristic is the guarantee for the achievement of the optimal ξ value in the search process.

In ***Algorithm 1***, the iterative search method is depicted, where $\xi_{\min}$, $\xi_{\max}$, ξ_{mid} are updated by comparing the calculated bits and the bit constraint R_c in each iteration. $\xi_{\min}$ can achieve the minimum bits, while $\xi_{\max}$ achieves the maximum bits. If the optimal ξ is not in the range of [$\xi_{\min}$, $\xi_{\max}$], the range is extended by the factors $e_{\min}$ and $e_{\max}$ for $\xi_{\min}$ and $\xi_{\max}$, respectively, which are both set as 10 in the experiment. The finally achieved ξ_{mid} is denoted as the best ξ value ξ_{best}.

After the best ξ is achieved, (26) and (27) can be used to get the bit allocation for the i-th CTU. It is then constrained by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

$$r_{i,opt} = \max\left(r_{i,\min}, \min\left(\max\left(r_{i,d}, r_i\right), r_{i,\max}\right)\right), \tag{33}$$

where $r_{i,\min}$ and $r_{i,\max}$ are the bit constraints from the previous frame with the same TL, $r_{i,opt}$ is the optimal bit allocation for the i-th CTU from the cooperative bargaining game.

Algorithm 1 Iterative Solution Search for Lagrange Multiplier ξ

Input:

The ξ range: $[\xi_{\min}, \xi_{\max}]$, and ξ_{mid}=mean($\xi_{\min}$, $\xi_{\max}$)

The current frame-level bit constraint for CTUs with 1.0-order and 1.5-order R-D models: R_c

Output:

The best Lagrange Multiplier: ξ_{best}

for Iter$\in$[1, Iter_max] **do**

  Calculate the consumed bits $g(\xi_{\min})$, $g(\xi_{\max})$, $g(\xi_{mid})$ for $\xi_{\min}$, $\xi_{\max}$, ξ_{mid};

  if $R_c<g(\xi_{mid})$ and $R_c>g(\xi_{\min})$

    $\xi_{\max}=\xi_{mid}$, $\xi_{\min}=\xi_{\min}$;

  else if $R_c<g(\xi_{\max})$ and $R_c>g(\xi_{mid})$

    $\xi_{\max}=\xi_{\max}$, $\xi_{\min}=\xi_{mid}$;

  else if $R_c<g(\xi_{\min})$

    $\xi_{\max}=\xi_{\min}$, $\xi_{\min}=\xi_{\min}/e_{\min}$;

  else if $R_c>g(\xi_{\max})$

    $\xi_{\min}=\xi_{\max}$, $\xi_{\max}=e_{\max}\,\xi_{\max}$;

  end if

  ξ_{mid}=mean($\xi_{\min}$, $\xi_{\max}$);

end for

Determine the achieved ξ_{mid} as ξ_{best}.

Motivated by the R-Q_{step} relationships in (17) and (18), an estimated bit budget $r_{i,est}$ can be given by considering the frame-level QP change

$$r_{i,est} = r_{i,ref} / \delta^{\eta}, \tag{34}$$

where δ is defined in (14), $r_{i,ref}$ is the consumed bits from the reference CTU, η is equal to 1.0 and 1.5 for the 1.0-order and 1.5-order R-D models, respectively. Then, $r_{i,\min}$ and $r_{i,\max}$ can be set as

$$r_{i,\min} = \left(1 - P_{\min}\right) r_{i,est}, \tag{35}$$

$$r_{i,\max} = \left(1 + P_{\max}\right) r_{i,est}, \tag{36}$$

where $P_{\min}$ and $P_{\max}$ are the allowable range percentages for the minimum and maximum bit constraints, respectively.

It should be noted that the above bit allocation scheme is for CTUs which select 1.0-order and 1.5-order R-D models. For SKIP_Most_CTUs, the bit budgets are the same as the collocated CTUs in the previous frame with the same TL.

D. CTU-Level Two-stage Refinement Based Bit Allocation and QP Determination

After the optimal Lagrange multiplier ξ is achieved by the proposed iterative search method, the optimal CTU-level bit allocation $R_{i,opt}$ can be calculated from (26) and (27), where parameter $C_{1,i}$ and $C_{2,i}$ can be achieved from the *C_Map*.

Before encoding the current CTU, we firstly get the bit weight BW_i for the i-th CTU

$$BW_i = r_{i,opt} \Big/ \sum_{j=1}^{N} r_{j,opt}. \tag{37}$$

A similar two-stage remaining bit refinement method in [7] can be used for the CTU-level bit allocation. The target remaining bits TRB_i for the i-th CTU is

$$TRB_i = \left(\sum_{k=i}^{N} BW_k \Big/ \sum_{j=1}^{N} BW_j \right) R_c, \tag{38}$$

where R_c is the target frame-level bit constraint for CTUs with 1.0-order and 1.5-order R-D models.

The target remaining bits are calculated by assuming all CTUs are encoded with the optimal bit consumption. However, in practice, such accurate bit consumption can not always be achieved. Therefore, the mismatch gap between the target remaining bits TRB_i and the actual remaining bits ARB_i are proposed to be compensated by using a smoothing window.

Then the refined remaining bits RRB_i is calculated as

$$RRB_i = ARB_i + \left(ARB_i - TRB_i\right) \times NRC / SWS, \tag{39}$$

where NRC is the number of remaining CTUs, and SWS is the size of CTU-level bit smoothing window. This smooth window will benefit the quality smoothness and bit accuracy.

The final optimal bit allocation $R_{i,opt}$ for i-th CTU can be achieved by RRB_i and bit weight BW_i

$$r_{i,opt} = \left(BW_i \Big/ \sum_{j=i}^{N} BW_j \right) RRB_i. \tag{40}$$

Based on the R-Q_{step} relationships in (17) and (18), we can get the Q_{step} for the two R-D models, respectively

$$Q_{step,i} = k_i C_{1,i} / R_i, \tag{41}$$

$$Q_{step,i} = k_i \left(C_{2,i} / R_i \right)^{2/3}, \tag{42}$$

where k_i is from the CTU-level D-Q model in (16), and $C_{1,i}$ and $C_{2,i}$ are obtained from the *C_Map*. Based on (40), (41) and (42), we can determine the Q_{step} and QP for the current CTU.

E. Overall Rate Control Algorithm Flow

The overall RC algorithm flow of the proposed MLGT method can be illustrated in Fig. 1. In addition, to enhance the RC performances, the following steps are also implemented.

More bit resources are expected to be allocated to inter frames to best exploit the potentials of the proposed joint MLGT framework to improve the R-D performances of inter frames. The intra frame QP is adjusted with an increment of 5 in the experiments when compared with the original HM16.8. For inter frame RC, adaptive hierarchical bit ratios are used to take into account the quality dependency for inter frames in the LB coding structure with GOP size=4. The adaptive bit ratios are set for different frames in a GOP according to the *bpp* values as in Table IV.

TABLE IV
BPP BASED FRAME-LEVEL ADAPTIVE BIT RATIOS FOR GOP FRAMES

Frame	bpp>0.25	0.12<bpp≤0.25	0.06<bpp≤0.1	bpp≤0.06
0	2	2	2	2
1	3	3	3	3
2	2	2	4	2
3	4	6	8	10

The above adjustments on intra and inter frames can let inter frames have more bit resources to make the proposed MLGT RC method better exploit the performance potentials, and can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

also make the inter frames have much smoother results on the consumed bits and coding quality.

According to the flow shown in Fig.1, after frame-level RC, the joint MLGT based CTU-level RC framework can be usd to improve the R-D modeling accuracy and RC performances.

V. Experimental Results

A. Experiment Setup

To verify the performances of the proposed MLGT based RC optimization method for HEVC, we have implemented it in the latest HM16.8 platform [17]. The original HM16.8 benchmark ("HM16.8-RLRC") [13]-[17] is also tested, which adopts the R-λ model based RC scheme. We have also implemented two other state-of-the-art RC schemes for comparison, including [8] and [10], which are denoted as "TIP16-Wang" and "TIP13-Seo", respectively. In addition, as a reference, the FixedQP method is also tested, which encodes the video sequences multiple times with different fixed QPs to achieve the optimal coding performance. FixedQP method [6]-[7] usually provides good R-D performances and quality smoothness results in most cases, but the disadvantages are the extremely long encoding time from the multiple attempts and the bit rate accuracy is usually not high. In addition, to demonstrate the effectiveness and necessity of using machine learning, the scene change based R-D model prediction method in Section II-B is also tested with the same game theory based bit allocation scheme. The scene change based and machine learning based methods are denoted as "Proposed SCGT" and "Proposed MLGT", respectively.

The coding configurations are set as follows: the coding structure is LB, the profile is main, CTU size is 64, CTU maximum depth is 4, intra period is 32, fast search is enabled, adaptive hierarchical bit allocation is enabled, CTU-level RC is open, CTU separate model is enabled, initial QP setting is from adaptive calculation. In the following sections, the results on R-D performance, quality smoothness, bit rate accuracy, coding complexity, buffer occupancy and subjective quality will be compared and analyzed for different RC methods.

B. Analysis for Intra Frame QP Increment

The intra QP increments of 0, 3, 5, 7 over the original HM16.8 are tested for the proposed MLGT RC scheme, which are denoted as AQP0, AQP3, AQP5 and AQP7, respectively. Compared with AQP5, the results of AQP0, AQP3 and AQP7 are listed in Table V. Bjøntegaard delta peak-signal-to-noise-ratio (BD-PSNR) (dB) and Bjøntegaard delta structural similarity (BD-SSIM) results are listed for R-D performance comparisons. The bit rate accuracy differences are denoted as DBRA (%). To measure the quality smoothness, the standard variances of PSNR (dB) and SSIM are calculated, respectively, and the corresponding differences are denoted as DSPSNR (dB) and DSSSIM, respectively. In Table V, AQP5 has the best results on R-D performances, bit rate accuracy and quality smoothness. Therefore, the intra QP increment of 5 is used in the proposed MLGT based RC scheme and for further sections.

Table V
Rate Control Performance Comparisons using Different Intra QP Increments (Compared with AQP5)

Class	AQP0					AQP3					AQP7				
	BD-PSNR	**BD-SSIM**	**DBRA**	**DSPSNR**	**DSSSIM**	**BD-PSNR**	**BD-SSIM**	**DBRA**	**DSPSNR**	**DSSSIM**	**BD-PSNR**	**BD-SSIM**	**DBRA**	**DSPSNR**	**DSSSIM**
A	-0.194	-0.0013	1.625	0.320	0.0022	0.028	0.0005	0.860	0.054	0.0003	-0.051	-0.0006	-0.637	0.077	0.0006
B	-0.401	-0.0011	-0.402	0.265	0.0009	-0.105	0.0000	0.274	0.084	0.0002	-0.072	-0.0007	-1.041	0.096	0.0007
C	-0.341	-0.0056	-1.301	0.518	0.0062	-0.061	-0.0013	0.056	0.196	0.0020	-0.077	-0.0007	-0.035	-0.067	0.0003
D	-0.312	-0.0042	-1.882	0.509	0.0053	-0.061	-0.0009	-0.526	0.221	0.0017	0.023	0.0002	-0.061	-0.128	-0.0005
E	-0.178	-0.0004	-4.999	0.592	0.0014	-0.008	-0.0003	-0.855	0.282	0.0010	-0.212	-0.0003	-0.443	0.287	0.0005
Average	**-0.285**	**-0.0025**	**-1.392**	**0.441**	**0.0032**	**-0.041**	**-0.0004**	**-0.038**	**0.167**	**0.0010**	**-0.078**	**-0.0004**	**-0.444**	**0.053**	**0.0003**

C. R-D Performance

Compared with HM16.8-RLRC, the BD-PSNR (dB) and BD-SSIM results on HM16.8 for all other RC methods are listed in Table VI, including TIP16-Wang, TIP13-Seo, the proposed SCGT, the proposed MLGT and FixedQP methods.

The proposed method can averagely achieve 0.334 dB BD-PSNR and 0.0027 BD-SSIM gains, respectively, which are very remarkable and approximate the R-D performance limits given by FixedQP, which are 0.366 dB BD-PSNR and 0.0030 BD-SSIM gains, respectively. As the other two state-of-the-art one-pass RC methods, TIP16-Wang and TIP13-Seo have much worse R-D performances than the benchmark HM16.8-RLRC method. The proposed MLGT method also outperforms the proposed SCGT method. It can be seen that compared with the scene change based method, the improved R-D model prediction accuracy by the proposed machine learning based method can help R-D performances. In summary, the proposed MLGT method can work better than all the other one-pass RC methods on R-D performances, and the achieved gains are very close to those provided by the multiple-pass FixedQP method.

In Table VI, compared with the benchmark HM16.8-RLRC, FixedQP provides the best R-D performances. FixedQP needs multiple encoding attempts using different QPs to achieve the closest bit rate to the target bit rate. The best R-D performance can be usually achieved by sacrificing much more encoding time than all the other one-pass RC methods. This drawback prevents FixedQP from many practical multimedia processing applications. The reasons for the R-D performance gains from the proposed MLGT method originate from three parts. First, the learned CTU-level R-D models can discriminatively handle each CTU. Second, the CTU-level bit allocation is optimized by the cooperative bargaining game to achieve better R-D performances. Lastly, intra frame QPs are adjusted

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

to make inter frames have more bit resources, and then the proposed MLGT bit allocation scheme can better exploit the potentials to improve RC performances.

D. Quality Smoothness

The quality consistency can be measured by the smoothness of frame-level PSNR and SSIM, which is a very important factor to evaluate the human visual experience on video quality. In this work, the standard variances of frame-level PSNR and SSIM are used to evaluate the quality smoothness, which are denoted as S_PSNR (dB) and S_SSIM, respectively.

In Table VII, the quality smoothness results are compared for different RC methods. Since FixedQP adopts the same QP for all frames, it achieves the minimal quality fluctuation results for almost all sequences. For the one-pass RC methods, TIP16-Wang has the maximum fluctuations, while the proposed MLGT method achieves the minimal fluctuations in terms of both S_PSNR and S_SSIM, which are 0.908 dB and 0.0061, respectively. The achieved reductions on quality fluctuations from the proposed MLGT method are very remarkable, and very close to those from FixedQP which are 0.559 dB and 0.0041 on S_PSNR and S_SSIM, respectively. Due to the inaccurate R-D modeling, the proposed SCGT method has larger quality fluctuations than the proposed MLGT method. In summary, the proposed MLGT method can have much lower quality fluctuations in terms of S_PSNR and S_SSIM than all other state-of-the-art one-pass RC methods, and the achieved quality fluctuation results are also very close to those from FixedQP.

TABLE VI
R-D PERFORMANCE COMPARISONS FOR DIFFERENT RC METHODS (COMPARED WITH HM16.8-RLRC)

Class	TIP16-Wang		TIP13-Seo		Proposed SCGT		Proposed MLGT		FixedQP	
	BD-PSNR	**BD-SSIM**	**BD-PSNR**	**BD-SSIM**	**BD-PSNR**	**BD-SSIM**	**BD-PSNR**	**BD-SSIM**	**BD-PSNR**	**BD-SSIM**
A	-0.414	-0.0079	0.203	0.0015	0.015	-0.0014	0.395	0.0024	0.490	0.0026
B	-0.612	-0.0065	-0.812	-0.0078	-0.069	-0.0017	0.417	0.0010	0.437	0.0017
C	-0.343	-0.0064	0.215	0.0036	-0.177	-0.0027	0.305	0.0056	0.253	0.0053
D	-0.073	-0.0014	0.175	0.0008	-0.204	-0.0041	0.341	0.0041	0.362	0.0051
E	0.353	0.0004	-0.238	-0.0002	0.870	0.0017	0.213	0.0005	0.290	0.0003
Average	**-0.218**	**-0.0043**	**-0.091**	**-0.0004**	**0.087**	**-0.0016**	**0.334**	**0.0027**	**0.366**	**0.0030**

TABLE VII
QUALITY FLUCTUATION COMPARISONS FOR DIFFERENT RC METHODS

Class	HM16.8		TIP16-Wang		TIP13-Seo		Proposed SCGT		Proposed MLGT		FixedQP	
	S_PSNR	**S_SSIM**	**S_PSNR**	**S_SSIM**	**S_PSNR**	**S_SSIM**	**S_PSNR**	**S_SSIM**	**S_PSNR**	**S_SSIM**	**S_PSNR**	**S_SSIM**
A	1.780	0.0114	2.213	0.0178	1.292	0.0078	1.732	0.0091	0.938	0.0061	0.560	0.0032
B	1.173	0.0050	2.134	0.0157	1.527	0.0090	1.509	0.0075	0.844	0.0037	0.647	0.0029
C	1.408	0.0161	2.136	0.0247	1.242	0.0133	1.702	0.0157	0.850	0.0088	0.629	0.0077
D	1.749	0.0162	1.708	0.0151	1.602	0.0152	1.859	0.0163	1.213	0.0109	0.638	0.0059
E	1.362	0.0022	1.037	0.0013	1.741	0.0026	1.663	0.0017	0.696	0.0008	0.320	0.0005
Average	**1.495**	**0.0102**	**1.846**	**0.0149**	**1.481**	**0.0096**	**1.693**	**0.0100**	**0.908**	**0.0061**	**0.559**	**0.0041**

The quality fluctuation reductions of the proposed MLGT method mainly originate from three reasons. First, the intra frame QP is effectively adjusted, which makes the following B frames have more bit resources for encoding and their allocated QP values become much closer to intra frame QP. Therefore, the quality differences between intra and inter frames become smaller. Second, adaptive bit ratios for frames in a GOP are set closer than those in HM16.8-RLRC to reduce quality fluctuations. Lastly, the minimum utility is defined by the reference coding distortion and frame-level QP change, which makes CTU-level bit allocation much more accurate, as well as frame-level. Therefore, the proposed minimum utility definition reduces the possibility of quality fluctuations from insufficient or excessive available bit resources.

E. Bit Rate Accuracy

Bit rate accuracy is one of the key optimization objectives in RC. The actual bit rates should be very close to the target bit rates. If the mismatch is large, this RC method is not preferred. The results on bit rate accuracy (%) are compared in Table VIII, in which it can be seen that the proposed MLGT method can achieve the highest bit rate achievement result in average, which is up to 98.734%. The other one-pass RC methods and FixedQP cannot get the desirable bit rate achievements.

TABLE VIII
BIT RATE ACCURACY COMPARISONS FOR DIFFERENT RC METHODS

Class	HM16.8	TIP16-Wang	TIP13-Seo	Proposed SCGT	Proposed MLGT	FixedQP
A	98.843	89.390	99.832	95.667	97.851	98.164
B	97.478	94.990	96.610	95.788	96.931	97.056
C	98.488	89.227	99.292	97.960	99.825	96.124
D	97.441	77.631	99.204	97.632	99.630	96.833
E	93.170	96.324	64.402	84.898	99.433	97.915
Average	**97.084**	**89.512**	**91.868**	**94.389**	**98.734**	**97.218**

F. Computational Complexity

The computational complexity can be roughly measured by encoding time. For fairness and accuracy, results are collected from separate simulations on a personal computer with Intel (R) Core TM i3-2330M CPU at 2.2GHz and 6G RAM memory.

The encoding complexity results for different RC methods are compared in Table IX, where the complexity ratios are achieved by comparing with the benchmark HM16.8-RLRC. It can be seen that, TIP16-Wang and TIP13-Seo have 3.68% and 0.58% increases, respectively, while the proposed MLGT has 5.01% increase. In general, all one-pass RC methods have the same level coding complexity results and the complexity increases can be negligible in practical encoding.

TABLE IX
ENCODING COMPLEXITY COMPARISONS FOR DIFFERENT RC METHODS

Class	HM16.8-RLRC	TIP16-Wang	TIP13-Seo	Proposed MLGT
A	100.00	101.49	100.72	104.04
B	100.00	98.68	100.29	104.02
C	100.00	106.64	100.87	106.44
D	100.00	109.93	100.71	105.82
E	100.00	101.65	100.30	104.75
Average	**100.00**	**103.68**	**100.58**	**105.01**

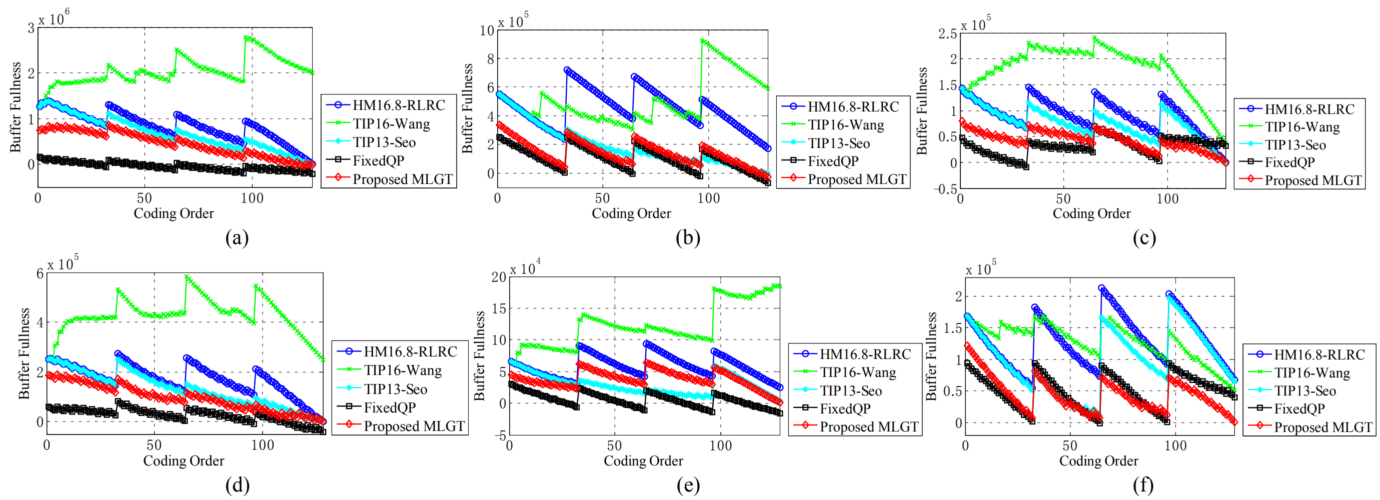
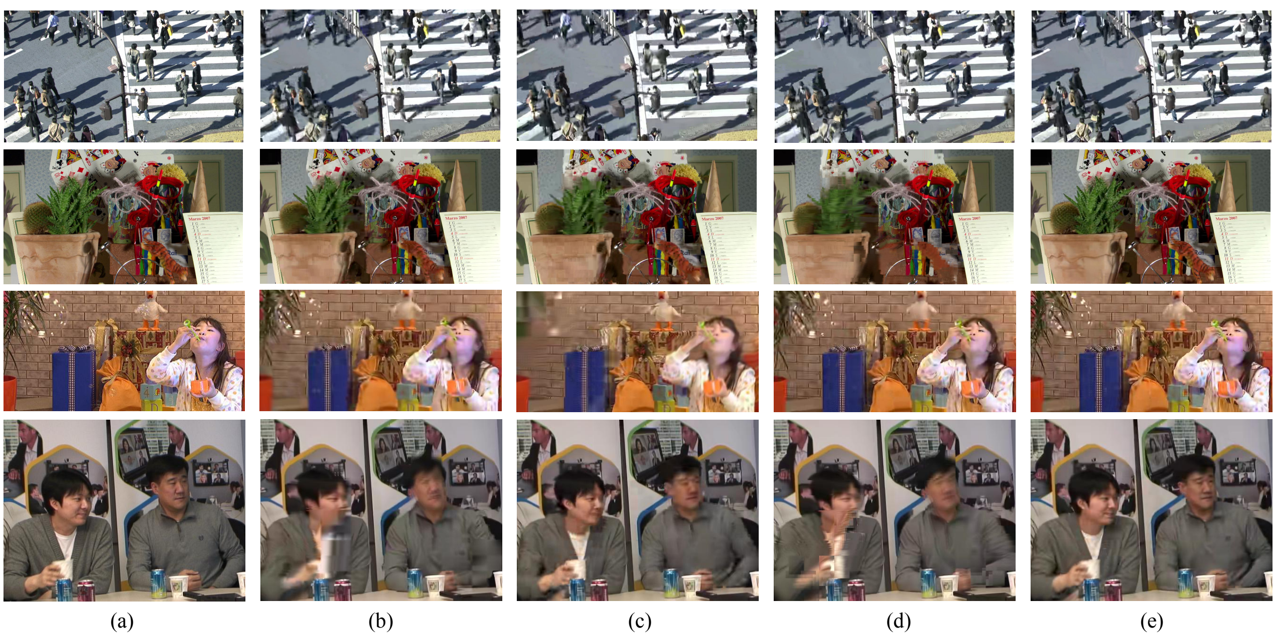


Fig. 4. Buffer fullness comparisons. (a) PeopleOnStreet (target bitrate: 1400 kbps); (b) BQTerrace (target bitrate: 700 kbps) (c) BasketballDrill (target bitrate: 350 kbps); (d) RaceHorsesC (target bitrate: 350 kbps); (e) BQSquare (target bitrate: 125 kbps); (f) FourPeople (target bitrate: 300 kbps).



(a) (b) (c) (d) (e)

Fig. 5. Subjective visual fidelity comparisons. (a) Original frame; (b) HM16.8-RLRC; (c) TIP16-Wang; (d) TIP13-Seo; (e) Proposed MLGT. Exemplary results are listed from top to bottom: PeopleOnStreet (target bitrate: 1800 kbps); Cactus (target bitrate 900 kbps); PartyScene (target bitrate: 550 kbps); FourPeople (target bitrate: 350 kbps).

G. Buffer Occupancy

For practical video transmission, the generated bit streams will be transmitted on the communication channel, which has a bandwidth constraint for video streams. Because there always exists the mismatch between the real-time generated bits from video encoding and the bandwidth constraint, a fixed size buffer is used to temporally store the bits that cannot be transmitted immediately. Obviously, buffer overflow occurrences are highly related to both the RC method and the buffer size. Once overflow or underflow cases happen, the video transmission system will skip some frames. Due to the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2017.2745099, IEEE Transactions on Image Processing

information loss, the video decoding end will bring bad visual experiences. To this end, stable buffer occupancy control is a very important issue for RC performance evaluation.

In Fig. 4, the buffer fullness results in different RC methods are plotted and compared, where we can find that TIP16-Wang is the most likely to encounter more overflow cases. The other two RC methods, including HM16.8-RLRC and TIP13-Seo, also have much higher buffer occupancy results than the proposed MLGT method. A lower buffer footprint can be maintained by the proposed method for most time, which indicates that it will have much less overflow cases. For underflow cases, from Fig. 4, FixedQP has a lot of underflow cases while the other RC methods have no underflow cases. In summary, the proposed MLGT method obtains the best performances in buffer occupancy control to effectively avoid overflow and underflow cases, which is very helpful to achieve high decoding performances and comfortable visual experiences at the receiver end of video transmission system.

H. Subjective Quality Comparison

The subjective visual fidelity comparisons of different RC methods can be seen from the exemplary frames listed in Fig. 5. Compared with the proposed MLGT RC method, the other RC methods may much more easily have serious information loss, color artifacts, blocking effect, structural deformation, which significantly degrade the visual quality. Besides, as a matter of fact, another very influential factor to the observer's visual experience is the quality smoothness. As we know, the abrupt quality degradation and jitters will cause serious visual discomforts. From the above quality fluctuation comparisons in terms of S_PSNR and S_SSIM, the proposed MLGT method can have the minimal quality fluctuations. Moreover, from the test results of compressed bit streams and reconstructed videos in [42], we can find the proposed method can have much smoother subjective visual quality. In summary, the proposed MLGT RC method can work better than the other algorithms on subjective coding quality.

VI. Conclusion

The joint machine learning and game theory modeling (MLGT) based rate control optimization framework has been proposed for HEVC in this paper. Three main contributions have been presented. First, a machine learning based R-D model prediction method has been proposed to enhance the CTU-level R-D modeling accuracy. We also utilize the scene change detection for R-D model prediction, and its coding performances are compared with the machine learning based method. Second, a mixed R-D model based game theory approach has been proposed for bit allocation optimization, where the minimum utility definition is refined. After proving the convexity for mixed R-D model based mapping functions, NBS solution is achieved by the proposed iterative solution search method. A two-stage remaining bit refinement method is used for bit allocation and QP determination. Finally, the intra frame QP and adaptive bit ratios among frames are effectively refined to better exploit the potentials of the proposed MLGT method. Experimental results demonstrate that the proposed MLGT RC method can achieve much better results on R-D performance, bit rate accuracy, quality smoothness, buffer occupancy control and subjective visual experience than all other state-of-the-art one-pass RC methods, and the achieved results on R-D performances and quality smoothness are close to the performances from FixedQP.

Appendix A

In this appendix, we prove Theorem 1. Assume there are two feasible utility sets: $U_x=\{U_{x,i}(r_{x,i}),\ i=1, 2, ..., N\}$ and $U_y=\{U_{y,i}(r_{x,i}),\ i=1, 2, ..., N\}$. According to the definition for convex function [28], the mapping function f: $U\rightarrow r$ is convex when (43) is satisfied

$$\sum_{i=1}^{N} f\left(\theta U_{x,i}+(1-\theta)U_{y,i}\right)\le \sum_{i=1}^{N}\left[\theta f\left(U_{x,i}\right)+(1-\theta)f\left(U_{y,i}\right)\right], \quad (43)$$

where $0<\theta<1$, f can be f_1 in (8) or f_2 in (9) for different CTUs.

As we know, each CTU uses a separate R-D model and is independent of other CTUs. Therefore, if we can prove

$$f\left(\theta U_{x,i}+(1-\theta)U_{y,i}\right)\le \theta f\left(U_{x,i}\right)+(1-\theta)f\left(U_{y,i}\right), \quad (44)$$

then (43) can be proved.

If $U_{y,i}>U_{x,i}$, from the convex function definition in [28], (44) is also equivalent to

$$t\left(U_{x,i},U_{y,i},\theta\right)=f\left(U_{y,i}\right)-f\left(U_{x,i}\right)-f'\left(U_{x,i}\right)\left(U_{y,i}-U_{x,i}\right)\ge 0. \quad (45)$$

From (8) and (9), $f_1'(U_i)=C_{1,i}>0$ and $f_2'(U_i)=3C_{2,i}U_i^{1/2}/2>0$, then f_1 and f_2 are both monotonically increasing functions.

When the mapping function f_1 in (8) is applied

$$t\left(U_{x,i},U_{y,i},\theta\right)=C_{1,i}U_{y,i}-C_{1,i}U_{x,i}-C_{1,i}\left(U_{y,i}-U_{x,i}\right)=0. \quad (46)$$

Otherwise, the mapping function f_2 in (9) is applied

$$\begin{aligned} t\left(U_{x,i},U_{y,i},\theta\right)&=C_{2,i}U_{y,i}^{3/2}-C_{2,i}U_{x,i}^{3/2}-3C_{2,i}U_{x,i}^{1/2}\left(U_{y,i}-U_{x,i}\right)/2\\ &=C_{2,i}\left(U_{y,i}^{1/2}-U_{x,i}^{1/2}\right)\left(U_{y,i}+U_{y,i}^{1/2}U_{x,i}^{1/2}+U_{x,i}-3U_{x,i}^{1/2}\left(U_{y,i}^{1/2}+U_{x,i}^{1/2}\right)/2\right).\\ &=C_{2,i}\left(U_{y,i}^{1/2}-U_{x,i}^{1/2}\right)^2\left(U_{y,i}^{1/2}+U_{x,i}^{1/2}/2\right)\ge 0 \end{aligned} \quad (47)$$

Obviously, whether a CTU selects the mapping function f_1 or f_2, the convexity can be proved separately, then (44) can be proved for every CTU, and (43) can be proved for all CTUs in a frame. Therefore, the convexity of the utility mapping function in the feasible utility set U can be proved for the mixed R-D model based cooperative bargaining game.

This completes the proof of Theorem 1.

Appendix B

In this appendix, we will discuss the cubic function solution in (27) and the monotonic relationship between r_i and ξ.

From (27), we can have

$$\begin{aligned} \frac{2}{3\xi}&=C_{2,i}^{2/3}U_{i,d}r_i^{1/3}-r_i=-r_i^{1/3}\left(r_i^{1/3}-C_{2,i}^{1/3}U_{i,d}^{1/2}\right)\left(r_i^{1/3}+C_{2,i}^{1/3}U_{i,d}^{1/2}\right).\\ &=-r_i^{1/3}\left(r_i^{1/3}-r_{i,d}^{1/3}\right)\left(r_i^{1/3}+C_{2,i}^{1/3}U_{i,d}^{1/2}\right)<0 \end{aligned} \quad (48)$$

Therefore, ξ<0. Then, we denote the following x, S and T

$$x = r_i^{1/3}, \tag{49}$$

$$S = -C_{2,i}^{2/3} U_{i,d} = -r_{i,d}^{2/3}, \tag{50}$$

$$T = 2/(3\xi), \tag{51}$$

where x>0, S<0, and T<0. The cubic function in (27) becomes

$$x^3 + Sx + T = 0. \tag{52}$$

We denote a function $p(x)$ as

$$p(x) = x^3 + Sx = x\left(x - \sqrt{-S}\right)\left(x + \sqrt{-S}\right). \tag{53}$$

The derivative of $p(x)$ also can be calculated as

$$p'(x) = 3x^2 + S = \left(\sqrt{3}x - \sqrt{-S}\right)\left(\sqrt{3}x + \sqrt{-S}\right). \tag{54}$$

If $x > \sqrt{-S/3}$, p and x exhibit a monotonically increasing relationship. As $p(x)$=-T>0, from (53) and x>0, we can have $x > \sqrt{-S} > \sqrt{-S/3}$. Therefore, in the range of all available x values, p and x exhibit a monotonically increasing relationship. If ξ increases, $p(x)$=-T=-2/(3ξ) also increases. Then ξ and p exhibit a monotonically increasing relationship. Therefore, ξ and x exhibit a monotonically increasing relationship in the available range of $x > \sqrt{-S} = r_{i,d}^{1/3}$.

Next, we get the unique solution of x. The discrimination function Ω is denoted as

$$\Omega = (S/3)^3 + (T/2)^2 = \left(1/\xi^2 - r_{i,d}^2/3\right)/9. \tag{55}$$

According to Cardano's method [30], the solutions for x are

$$\begin{cases} x_1 = \sqrt[3]{-T/2 + \sqrt{\Omega}} + \sqrt[3]{-T/2 - \sqrt{\Omega}} \\ x_2 = \omega \sqrt[3]{-T/2 + \sqrt{\Omega}} + \omega^2 \sqrt[3]{-T/2 - \sqrt{\Omega}} \\ x_3 = \omega^2 \sqrt[3]{-T/2 + \sqrt{\Omega}} + \omega \sqrt[3]{-T/2 - \sqrt{\Omega}} \end{cases}, \tag{56}$$

where $\omega = \left(-1 + \sqrt{3}i\right)/2$.

If Ω>0, the cubic function has a real root x_1 and two complex roots x_2 and x_3. If Ω=0, it has three real roots. Since S, T≠0, then $x_1 \neq x_2 = x_3$. If Ω<0, it has three different real roots. From video coding practice, x>0 and x is unique for a fixed ξ. Let Ω≥0, from (55) we have

$$-\sqrt{3}/r_{i,d} \le \xi < 0, \tag{57}$$

From (54), when $x = -\sqrt{-S/3} = -r_{i,d}^{1/3}/\sqrt{3}$, we can have the local maximum of $p(x)$ in the range of x<0,

$$p\left(x = -\sqrt{-S/3}\right) = -\frac{2}{3} S \sqrt{-S/3} = 2r_{i,d}/3^{3/2}. \tag{58}$$

Let (58) be equal to -T, we can have the ξ at this local maximum point of $p(x)$

$$\xi\left(x = -\sqrt{-S/3}\right) = -\sqrt{3}/r_{i,d}. \tag{59}$$

From (55), it can be seen that ξ in (59) makes Ω=0. According to (56),

$$\begin{cases} x_1 = 2\sqrt[3]{-T/2} = 2r_{i,d}^{1/3}/\sqrt{3} \\ x_2 = x_3 = -\sqrt{-S/3} = -r_{i,d}^{1/3}/\sqrt{3} \end{cases}. \tag{60}$$

As aforementioned, x>0. In the range of x>0, only x=x_1 in (60) can make Ω=0. Based on the definition of x in (49) and bargaining game theory, we can know when $r_{i,d}<r_i<(8/3^{3/2})r_{i,d}$ (i.e. $(3^{3/2}/8)r_i<r_{i,d}<r_i$), Ω<0. For simplicity, $8/3^{3/2}\approx 1.54$, $3^{3/2}/8 \approx 0.65$. Moreover, when $r_i>1.54r_{i,d}$ (i.e. $r_{i,d}<0.65r_i$), Ω>0.

From the minimum utility definition in (13)

$$U_{i,d} = \gamma/(\delta D_{i,pre}) \approx \gamma/D_{i,cur} = \gamma U_i, \tag{61}$$

where γ is a constant 0.5, $D_{i,cur}$ is the coding distortion for the current i-th CTU. From the mapping function f_2 in (9)

$$\frac{r_{i,d}}{r_i} = \frac{C_{2,i}U_{i,d}^{3/2}}{C_{2,i}U_i^{3/2}} = \left(\frac{U_{i,d}}{U_i}\right)^{3/2} \approx \gamma^{3/2} = \left(\frac{1}{2}\right)^{3/2} \approx 0.35, \tag{62}$$

where we use $U_{i,pre,adjust}$ in (13) as the estimation of U_i. Then we can have $r_{i,d}\approx 0.35r_i<0.65r_i$ (i.e. $r_i\approx 2.86r_{i,d}>1.54r_{i,d}$), and Ω>0. Therefore, the solution of x can be expressed by x_1 in (56). r_i is achieved by $r_i = x_1^3$. From the above discussion, the range of r_i is $r_i>1.54r_{i,d}$. From (57), the range of ξ is $-\sqrt{3}/r_{i,d} < \xi < 0$.

In summary, based on the definition of disagreement point in the bargaining game theory, we can prove that the solution of r_i can be achieved for CTUs with 1.5-order R-D model, and a monotonically increasing relationship can be seen between ξ and r_i in (27) in the range of $r_i>r_{i,d}$.

REFERENCES

[1] G. J. Sullivan, J. R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[2] B. Bross, W.-J. Han, J. R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand, "High Efficiency Video Coding (HEVC) test specification draft 10", JCTVC-L1003, 12th JCTVC Meeting, Geneva, CH, Jan. 2013.

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

[4] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-Including High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1668-1683, Dec. 2012.

[5] S. Wang, S. Ma, S. Wang, D. Zhao, W. Gao, "Rate-GOP based rate control for High Efficiency Video Coding," *IEEE J. Selected Topics Signal Process.*, vol. 7, no. 6, pp. 1101-1111, Dec. 2013.

[6] W. Gao, S. Kwong, H. Yuan and X. Wang, "DCT coefficient distribution modeling and quality dependency analysis based frame-level bit allocation for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 139-153, Jan. 2016.

[7] W. Gao, S. Kwong, Y. Zhou and H. Yuan, "SSIM-based game theory approach for rate-distortion optimized intra frame CTU-level bit allocation," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 988-999, Jun. 2016.

[8] M. Wang, K. N. Ngan and H. Li, "Low-delay rate control for consistent quality using distortion-based Lagrange multiplier," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 2943-2955, July 2016.

[9] L. Xu, D. Zhao, X. Ji, L. Deng, S. Kwong, and W. Gao, "Window-level rate control for smooth picture quality and smooth buffer occupancy," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 723-734, Mar. 2011.

[10] C. W. Seo, J. H. Moon and J. K. Han, "Rate control for consistent objective quality in High Efficiency Video Coding," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2442-2454, June 2013.

[11] H. Choi, J. Yoo, J. Nam, J. Yoo, D. Sim, I. V. Bajić, "Pixel-wise Unified Rate-Quantization model for multi-level rate control," *IEEE IEEE J. Selected Topics Signal Process.*, vol. 7, no. 6, pp. 1112-1123, Dec. 2013.

[12] H. Choi, J. Nam, J. Yoo, D. Sim, I. V. Bajić, "Rate control based on unified RQ model for HEVC," JCTVC-H0213, 8-th JCTVC meeting, San Jose, CA, USA, Feb. 2012.

[13] B. Li, H. Li, L. Li, J. Zhang, "λ domain rate control algorithm for High Efficiency Video Coding," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3841-3854, Sept. 2014.

[14] B. Li, H. Li, L. Li, J. Zhang, "Rate control by R-lambda model for HEVC," JCTVC-K0103, 11th Meeting, Shanghai, CN, Oct. 2012.

[15] B. Li, H. Li, L. Li, "Adaptive bit allocation for R-lambda model rate control in HM," JCTVC-M0036, 13th JCTVC meeting, Incheon, KR, Apr. 2013.

[16] B. Li, D. Zhang, H. Li, J. Xu, "QP determination by lambda value," JCTVC-I0426, 9-th JCTVC meeting, Geneva, CH, May 2012.

[17] HM Reference Software 16.8. (2016, August) [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.8.

[18] Y. Guo, B. Li, S. Sun and J. Xu, "Rate control for screen content coding in HEVC," *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, 2015, pp. 1118-1121.

[19] F. Duanmu, Z. Ma and Y. Wang, "Fast mode and partition decision using machine learning for intra-frame coding in HEVC screen content coding extension," *IEEE J. Emerging Selected Topics Circuits Syst.*, vol. 6, no. 4, pp. 517-531, Dec. 2016.

[20] S. Sanz-Rodriguez and F. Diaz-de-Maria, "RBF-Based QP estimation model for VBR control in H.264/SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 9, pp. 1263-1277, Sept. 2011.

[21] J. Luo, I. Ahmad, and Y. Sun, "Controlling the bit rate of multi-object videos with noncooperative game theory," *IEEE Trans. Multimedia*, vol. 12, no. 2, pp. 97-107, Feb. 2010.

[22] I. Ahmad and Jiancong Luo, "On using game theory to optimize the rate control in video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 209-219, Feb. 2006.

[23] J. Nash, "The bargaining problem," *Econometrica*, vol. 18, no. 2, pp. 155-162, Apr. 1950.

[24] H. Yuan, S. Kwong, X. Wang, W. Gao and Y. Zhang, "Rate distortion optimized inter-view frame level bit allocation method for MV-HEVC," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2134-2146, Dec. 2015.

[25] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011. (2016, August) [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[26] B. Scholkopf, A. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Computation*, vol. 12, no. 5, pp. 1207-1245, May 2000.

[27] Chih-Wei Hsu and Chih-Jen Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. on Neural Networks*, vol. 13, no. 2, pp. 415-425, Mar 2002.

[28] Convex function. (2016, August) [Online]. Available: http://en.wiki pedia.org/wiki/Convex_function.

[29] Karush-Kuhn-Tucker conditions. (2016, August) [Online]. Available: http://en.wikipedia.org/wiki/Karush-Kuhn-Tucker_conditions.

[30] Cubic function. (2016, August) [Online]. Available: https://en.wiki-pedia.org/wiki/Cubic_function .

[31] L. Li, B. Li, H. Li, and C. W. Chen, "λ domain optimal bit allocation algorithm for High Efficiency Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol.PP, no.99, pp.1-1.

[32] G. Bjøntegaard, *Calculation of average PSNR differences between RD curves*, document VCEG-M33, Austin, TX, USA, Apr. 2001.

[33] Non-cooperative game theory. (2017, February) [Online]. https://en. wikipedia.org/wiki/Non-cooperative_game_theory.

[34] Cooperative game theory. (2017, February) [Online]. https://en. wikipedia.org/wiki/Cooperative_game_theory.

[35] Nash equilibrium. (2017, February) [Online]. Available: https://en. wikipedia.org/wiki/Nash_equilibrium.

[36] J. Lee, I. Shin and H. Park, "Adaptive intra-frame assignment and bit-rate estimation for variable GOP length in H.264," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 10, pp. 1271-1279, Oct. 2006.

[37] J. R. Ding and J. F. Yang, "Adaptive group-of-pictures and scene change detection methods based on existing H.264 advanced video coding information," *IET Image Processing*, vol. 2, no. 2, pp. 85-94, April 2008.

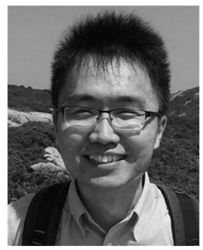
[38] I. Zupancic, M. Naccari, M. Mrak, E. Izquierdo, "Two-pass rate control for improved quality of experience in UHDTV delivery," *IEEE J. Selected Topics Signal Process.*, vol.PP, no.99, pp.1-13.

[39] S. Wang, A. Rehman, K. Zeng, J. Wang, Z. Wang, "SSIM-motivated two-pass VBR coding for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol.PP, no.99, pp.1-14.

[40] W. Gao and S. Kwong, "Phase congruency based edge saliency detection and rate control for perceptual image and video coding," *IEEE International Conference on Systems, Man, and Cybernetics*, Budapest, Hungary, 2016, pp. 264-269.

[41] W. Gao, S. Kwong, Y. Zhou, Y. Jia, J. Zhang and W. Wu, "Multiscale phase congruency analysis for image edge visual saliency detection," *International Conference on Machine Learning and Cybernetics*, Jeju, South Korea, 2016, pp. 75-80.

[42] Additional test results for the proposed MLGT rate control algorithm. (2017, February) [Online]. Available: https://drive.google.com/drive/ folders/0Bx6zfRE76vDCWFZYay1CYnJUSWM.



Wei Gao (S'16-M'17) received the Ph.D. degree in computer science from City University of Hong Kong, Kowloon, Hong Kong. He is currently an Assistant Professor with the National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University, Shenzhen, China. From 2012 to 2013, he was a Camera ISP Engineer with the OmniVision Technologies, Shanghai, China. In 2016, he was a Visiting Scholar with the Electrical Engineering Department, University of California, Los Angeles (UCLA), CA, USA. In 2017, he was a Postdoctoral Fellow with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, and a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, respectively. His research interests include image and video coding, perceptual image processing, multimedia signal processing, rate control, multimedia communication, machine learning and optimization.

Sam Kwong (M'93-SM'04-F'13) received the B.Sc. and M.Sc. degrees in electrical engineering from State University of New York, Buffalo, NY, USA, and the University of Waterloo, Waterloo, ON, Canada, in 1983 and 1985, respectively, and the Ph.D. degree from University of Hagen, North Rhine-Westphalia, Germany, in 1996. From 1985 to 1987, he was a Diagnostic Engineer with the Control Data Canada, Mississauga, ON, Canada. He later joined Bell Northern Research Canada, Ottawa, ON, Canada, as a Member of Scientific Staff, and the City University of Hong Kong (CityU), Hong Kong, as a Lecturer with the Department of Electronic Engineering in 1990. He is currently a Professor with the Department of Computer Science, CityU. His research interests include video coding, pattern recognition, and evolutionary algorithms. Prof. Kwong serves as an Associate Editor of the IEEE Transactions on Evolutionary Computation, the IEEE Transactions on Industrial Electronics and IEEE Transactions on Industrial Informatics, and the Journal of Information Science. He was elevated to IEEE Fellow for his contributions on Optimization Techniques for Cybernetics and Video Coding in 2014. He is also appointed as an IEEE Distinguished Lecturer for IEEE SMC society in 2017. He is currently the Vice President of Conferences and Meetings with IEEE Systems, Man and Cybernetics.

Yuheng Jia received the B.S. degree in automation and M.S. degree in control theory and control engineering from Zhengzhou University, Zhengzhou, China, in 2012 and 2015, respectively. He is currently working toward the Ph.D. degree in computer science at the City University of Hong Kong, Kowloon, Hong Kong. His research interests include machine learning, convex optimization, and sparse Bayesian learning.