A Novel Joint Rate Allocation Scheme of Multiple Streams

Hongfei Fan, Lin Ding, Huizhu Jia*, Xiaodong Xie

Abstract—Encoding multiple videos in parallel and transmitting them as one joint stream over a limited bandwidth have become a popular strategy for broadcasting, which brings an opportunity to allocate different bitrate for each sequence to meet different demands. In this paper, considering visual experience for human beings, we propose a joint rate allocation scheme aims to reach an equal visual quality among all sequences by minimizing the distortion variance of all the sequences (denoted as *minVAR* problems). Existing methods assigned bits directly in proportion to their complexity measures and we named them as complexity based allocation scheme (CAS) methods. CAS methods rely on the accuracy of the complexity measures which can hardly be improved under limited computing resources. Also complexities may not be directly related to the distortions. To address these problems, we present a novel joint rate-distortion (R-D) based allocation scheme (RDAS) in this paper. Our proposed scheme can fit for different R-D models and in our method we model the R-D relationship with a Hyperbolic function (RDAS-H). We also derive a closed-form solution of RDAS-H by a proposed joint R-D relationship. We integrated the RDAS-H method in HEVC reference software HM16.0. Experimental results demonstrate that our RDAS-H saves 75.29% variance on average over the related CAS-based method in [12], where we apply both LowDelay and RandomAccess configurations with four different overall bandwidths for all classes recommended by JCT-VC. Besides, RDAS-H also saves 36.62% variance on average over our previous method [17]. The proposed RDAS-H method improves the performance significantly while requiring negligible computational cost.

Index Terms—Joint rate allocation, Statistical multiplexing, Rate-Distortion model, *minVAR* problems, HEVC

I. INTRODUCTION

WITH the advancement in multimedia technology and digital communications, the amount of video programs has increased in terms of its applications. In this case a

H. Fan, L. Ding, H. Jia, and X. Xie are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: hffan@pku.edu.cn; lin.ding@pku.edu.cn; hzjia@pku.edu.cn; donxie@pku.edu.cn).

* the corresponding author, H. Jia is with Peking University, also with Cooperative Medianet Innovation Center and Beida (Binhai) Information Research.

single-program channel is substituted by broadcasting multiple videos over a limited bandwidth, where different videos are encoded in parallel and the bitrate is multiplexed to different videos to share the available bandwidth. To address such a joint rate allocation issue, a simple and direct solution is to encode each sequence with equal bitrate. Nevertheless, the difference in sequence content can be large and this scheme may lead to a significant difference in quality among the sequences. One sequence with high distortion will inevitably lead to poor visual experience for those who are watching this very sequence. Instead, an alternative solution is to allocate different bitrate for each sequence to meet demands like getting similar visual experience for the sequences. Specifically, statistical multiplexing is a way to dynamically allocate bitrate according to the different characteristics of video streams.

Typically, statistical multiplexing methods can be classified into two types according to different optimization targets, which are minimum average distortion (denoted as *minAVE* problem [1-5]) and minimum distortion variance (denoted as *minVAR* problem [6-17]). The group of frames (GOP) of all the sequences encoded in a fixed time interval is called a super GOP in this paper. The bitrates of different sequences are allocated by statistical multiplexing methods for each super GOP. The problem of statistical multiplexing can be formulated as to minimize the average distortion or the variance of distortion of all sequences, subject to the constraint that the sum of bitrates in one super GOP is limited by an overall bandwidth, and the channel should be fully utilized, i.e. the total quality should be maximized,

$$\begin{cases} \min_{R_k^i}\{ \bar{D}_k \}, & minAVE \ problem \\ \min_{R_k^i}\left\{ \frac{1}{N}\sum_{i=1}^{N}(D_k^i - \bar{D}_k)^2 \right\}, & minVAR \ problem \end{cases},$$

$$s.t. (R_c - \epsilon) \le \sum_{i=1}^{N} R_k^i \le R_c \qquad (1)$$

where R_k^i and D_k^i denote the bitrate and distortion for the i_{th} stream in the k_{th} super GOP, respectively; $\bar{D}_k$ is the average distortion of all sequences in the k_{th} super GOP. N is the total number of videos in the channel. R_c denotes the overall bandwidth for one super GOP. ϵ is a predefined small value according to different application. In this paper we consider an ideal situation where $\epsilon = 0$.

The objective of *minAVE* problems is to achieve maximum

average visual quality among the multiplexed videos. Although *minAVE* methods will improve the overall visual quality, it may result in different visual qualities among multiple sequences. In this case, one sequence with high distortion will lead to a poor visual experience for those who are watching this very sequence. Different from *minAVE* methods, *minVAR* methods aim to reach an equal visual quality among all sequences, which result in a better visual experience. In many practical applications such as broadcasting, *minVAR* method is a more preferred solution for a better overall visual experience. In this paper, we focus on *minVAR* problems.

In current methods to achieve *minVAR*, bits are assigned directly in proportion to their complexity measures. Since the bitrate allocation only depends on complexity measures, we classified them as complexity-based allocation scheme (CAS) methods. Generally in CAS method [6-15], a more accurate complexity measure can lead to a better performance by minimizing variance of multiple programs. According to the experimental results in [12] and [13], the performances of their proposed complexity measures are better than others. In [12], both the frame activity and the motion activity are used to characterize the video complexity. The weighting factor, which will be dynamically updated using the data from the previous GOP, is chosen to balance between frame activity and motion activity. In [13], inspired by Structural Similarity Index (SSIM), a new similarity index (SMI) is introduced to measure the similarity between two adjacent frames as temporal complexity. The authors also mentioned a spatial complexity for still image. Different from CAS, [16] proposed a scheme by searching a best distortion value for the frames to be allocated. An iterative golden-section search and a refinement search were carried out to reduce the computational cost. However, the performance of allocation will depend on the search quality.

Typically, the computing resource for look-ahead approach is limited which is far less than that for the formal encoding path. Although the accuracy of complexity measures has been improved in the state-of-art works, the efficiency of allocation scheme on *minVAR* problems can hardly be improved significantly under such limited computing constraints. Besides, there is no mathematical derivation to prove that assigning bits directly by the proportion of complexity measures can minimize the variance.

In our previous method [17], we proposed an allocation scheme by analyze R-D relationship and we use an inverse proportion function as a coarse estimation of the R-D relationship to simplify the derivation of the allocation formula. Results demonstrate that our previous method performs much better than CAS methods. Though [17] is good, the miscalculation will be amplified when the difference between the real distortion fed back by the encoder and the predicted distortion for the next super GOP is big. In this paper, we propose a novel joint rate-distortion based allocation scheme. Bandwidth is allocated considering not only complexity measures but also the R-D relationship built by the information fed back from the encoder. The allocation scheme is derived to meet the target function of *minVAR* problem mathematically. We can allocate the bandwidth properly if we can build the R-D

relationship for all sequences and the proposed scheme can work with all different R-D models. In the first step, the bitrate and distortion values of the newly coded super GOPs are obtained from the encoder. Then, the complexity measures of the next super GOPs are calculated by look-ahead approach. After that, bits are assigned to each sequence according to the allocation formula, considering the R-D model built by the feedback bitrate, distortion, and look-ahead complexity measure together. Finally, different sequences in the current super GOP are coded with the allocated bitrate. The contribution of this work mainly lies in the following two aspects:

1) We derive an R-D model based joint rate allocation scheme (RDAS) with the target function of *minVAR* problem, which is applicable to all kinds of R-D models. Furthermore, a Hyperbolic function based R-D model is applied to the RDAS scheme (RDAS-H) to obtain a better allocation result.

2) We derive a closed-form solution of RDAS-H instead of a brute-force solution to allocate the bitrate in O(1) time. To achieve this, we derive a joint R-D model which describes the relationship between the average rate and the distortion for a joint stream of multiple sequences based on information theory.

The rest of this paper is organized as follows. Section II gives the problem statement of statistical multiplexing. Section III gives the theoretical details of the proposed method. Section IV shows the experimental results to demonstrate the effectiveness of the proposed RDAS-H in comparison with that of [17] and the related CAS methods in [12]. This paper is finally concluded in section V.

II. JOINT RATE ALLOCATION PROBLEM AND RELATED WORK

In this section, we firstly state the *minVAR* problem in Part A. Then we briefly introduce the related work [12] in Part B which is a representative CAS method.

A. Problem Statement

In our system, multiple sequences are encoded as a joint stream and bitrate is allocated for each sequence at the beginning of each super GOP. Assume that the k_{th} super GOP is already encoded and we are now allocating the bitrate for the $(k+1)_{th}$ super GOP. The bitrate R_k^i and distortion D_k^i of the k_{th} super GOP for the i_{th} sequence can be obtained from the encoder. Then the *minVAR* problem can be rewritten as follows,

$$\min_{\hat{R}_{k+1}^i}\left\{\frac{\sum_{i=1}^{N}\left(D_{k+1}^i - \overline{D}_{k+1}\right)^2}{N}\right\}, \quad s.t.\,(R_c - \epsilon) \le \sum_{i=1}^{N}\hat{R}_{k+1}^i \le R_c \quad (2)$$

where $\overline{D}_{k+1}$ is the average distortion of all sequences in the $(k+1)_{th}$ super GOP. $\hat{R}_{k+1}^i$ is the allocated rate for the i_{th} sequence in the $(k+1)_{th}$ super GOP. In this paper, distortion is measured by mean square error (MSE) and bitrate is measured by bit per pixel (BPP).

According to related CAS methods, rates are allocated with the pre-process of information to meet the target function. Pre-process information is called complexity measure which is

obtained by the look-ahead approach. The complexity measure of the i_{th} video in the k_{th} super GOP is denoted as C_k^i. In this paper, we also use information of the last GOP fed back from the encoder. Thus, the bits are allocated as,

$$\hat{R}_{k+1}^i = \mathcal{R}\left(R_k^i, D_k^i, C_k^i, C_{k+1}^i\right) \qquad (3)$$

where the allocation scheme $\mathcal{R}(\cdot)$ is derived with the information meeting the constraints in Eq.2.

The framework of joint video coding is illustrated in Fig.1. Firstly, frames of individual sequences in one super GOP are pre-processed by look-ahead approach to get complexity values. Then, bitrate for each sequence is allocated by the Joint Rate Allocator using the allocation scheme in Eq.3. After that, Rate Controller will calculate QP and λ for individual sequence according to the allocated bitrate. Next, sequences are encoded separately by the encoder with the given QP and λ. At last, the encoded streams are sent to the Joint Buffer. At the same time, the real bitrate, distortion value and other parameters are sent back to the Joint Rate Allocator on demand for the rate allocation of the next super GOP.

In this paper, we only focus on the Joint Rate Allocator and assume that all the other parts in the system including Preprocessor, Rate Controller and Encoder are the same as the existing methods. In Preprocessor, we apply the same complexity measure as [12], and the complexity measure will be briefly introduced in Part B. In Rate Controller, the λ-domain rate control algorithm in [23] is applied, in which a Hyperbolic R−λ model is derived from the Hyperbolic R-D model. Besides, since our proposed scheme also use Hyperbolic R-D model, parameters of the model in the Rate Controller will be reused in the Joint Rate Allocator, which is denoted as 'other parameters' in Fig.1.
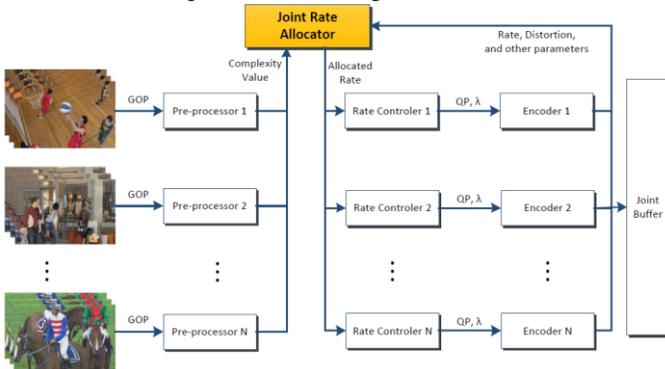

Fig.1. Framework of joint video coding.

B. Related Joint Rate Allocation Scheme in [12]

CAS methods to achieve $minVAR$ can be formulated as,

$$\hat{R}_k^i = \frac{C_k^i}{\sum_{i=1}^N C_k^i} \cdot R_c \qquad (4)$$

Since complexity measures are often computed with limited resource, a more precise complexity measure with negligible computational costs will improve the allocation efficiency.

Many researchers have proposed different methods for complexity measure computation. In [13] a novel complexity measure is proposed which outperforms that in [12] according to the results of [13]. However, [13] did not give all the constant values of their complexity measure and the performance of [12] and [13] are still comparable. So we apply the complexity measure in [12] in our system which is robust with little computational cost.

The complexity measure in [12] includes $C_{Texture}$ and C_{Motion} where $C_{Texture}$ denotes the complexity measure for the frame texture and C_{Motion} denotes that for motions. $C_{Texture}$ is calculated with,

$$C_{Texture} = \frac{1}{H \cdot W} \sum_{i=1}^{H-1} \sum_{j=1}^{W-1} (|lum(i,j) - lum(i+1,j)| + |lum(i,j) - lum(i,j+1)|) \qquad (5)$$

where H and W represent the height and the width of the frame. $lum(i,j)$ is the luminance value of pixel in the position (i,j) of the current frame. C_{Motion} is defined as,

$$C_{Motion} = \frac{1}{H \cdot W} \sum_{k=1}^{M-1} \sum_{i=1}^{H-1} \sum_{j=1}^{W-1} |lum(i,j,k) - lum(i,j,k+1)| \quad (6)$$

where $lum(i,j,k)$ is the luminance value of pixel at position (i,j) of the k_{th} frame and M is the number of frames in a super GOP. Combining Eq.5 and Eq.6 together,

$$C = \theta \cdot C_{Texture} + (1 - \theta)C_{Motion} \qquad (7)$$

where θ is a weighting factor chosen as 0 at the start of encoding. After encoding a super GOP, θ will be dynamically updated as the proportion of bits used for coding the I-frame of the whole GOP.

III. THE PROPOSED ALLOCATION SCHEME

In this section, we firstly derive an R-D model based allocation scheme in Part A, which is applicable to all kinds of R-D models. Then, in Part B, Hyperbolic R-D model, which outperforms other models on HEVC for most cases, is analyzed and applied to our allocation scheme. Besides, problems caused by applying Hyperbolic function are analyzed. Therefore, in Part C, we solve this problem and propose a joint R-D model and derive a closed-form solution for our scheme. And, in Part D, we give the detailed estimation of parameters. We finally analyze the R-D performance when applying our proposed allocation scheme in Part E.

A. Rate-Distortion Model Based Allocation Scheme (RDAS)

To allocate rate meeting the constraint in Eq.2 precisely, we take the Rate-Distortion relationship into consideration. Assume the relationship between rate and distortion as $R = \mathcal{R}(D)$. The rate of the $(k+1)_{th}$ super GOP of the i_{th} video can be expressed as,

$$\hat{R}_{k+1}^i = \mathcal{R}_{k+1}^i(\hat{D}_{k+1}^i) \qquad (8)$$

where $\hat{D}_{k+1}^i$ is the distortion corresponding to $\hat{R}_{k+1}^i$. $\mathcal{R}_{k+1}^i(\cdot)$ is the R-D relationship of all frames in the i_{th} video in the $(k+1)_{th}$ super GOP. Considering the target function of $minVAR$ problem given in Eq.2, the ideal allocation for the $(k+1)_{th}$ super GOP is as follows,

$$\forall i \in [1,N], \hat{D}_{k+1}^i = \hat{D}_{k+1}, \quad s.t. \sum_{i=1}^N \hat{R}_{k+1}^i = R_c \qquad (9)$$

which means the distortion values of every sequence in the $(k+1)_{th}$ super GOP are the same. Thus, we denote $\hat{D}_{k+1}$ as the corresponding distortion for an ideal allocation.

We sum up $\hat{R}_{k+1}^i$ of different sequences in the $(k+1)_{th}$ super GOP together,

$$R_c = \sum_{i=1}^N \hat{R}_{k+1}^i = \sum_{i=1}^N \mathcal{R}_{k+1}^i(\hat{D}_{k+1}) \qquad (10)$$

Assume that $\hat{R}_{k+1}^i > 0$, we divide Eq.8 by Eq.10 and get,

$$\hat{R}_{k+1}^i = \frac{\mathcal{R}_{k+1}^i(\hat{D}_{k+1})}{\sum_{i=1}^N \mathcal{R}_{k+1}^i(\hat{D}_{k+1})} \cdot R_c \qquad (11)$$

As we can see, Eq.11 is an R-D model based joint allocation scheme derived from the target function of $minVAR$ problem. The allocation scheme is applicable for all kinds of R-D models. Given the R-D relationship and $\hat{D}_{k+1}$, the allocation bitrate for the $(k+1)_{th}$ super GOP can be calculated directly.

B. The Proposed RDAS-H

In this part, we firstly review our previous work [17] in which we proposed a RDAS scheme which used an inverse proportion function for a coarse estimation of the R-D relationship (RDAS-I). To further improve the efficiency of RDAS-I, we propose a refined RDAS scheme in which the R-D relationship is modeled by the Hyperbolic function (RDAS-H).

In RDAS-I, a quadratic R-D relationship is modeled by expanding rate-distortion function into a Taylor series in [18]. Therefore, the inverse proportion function based R-D model is a coarse estimation of the quadratic model which is derived in [17] and [34] as,

$$D = \frac{\gamma \cdot C^2}{R} \qquad (12)$$

where γ is a constant parameter whose value reflects the characteristic of a sequence. C is the complexity measure of the sequence which can be calculated with Mean Absolute Difference (MAD) between the pixels of the current frame and that of the reference frame according to the derivation in [18]. In [17], the model in Eq.12 was verified to be acceptable when

the distortion value changes in a small range. When the k_{th} super GOP is encoded, we have,

$$R_k^i = \frac{\gamma_k^i \cdot C_k^{i\,2}}{D_k^i} \qquad (13)$$

We assume that the frame types are the same in each super GOP with both LowDelay and RandomAccess configurations and the characteristics between adjacent super GOPs of a sequence are similar. So,

$$\hat{\gamma}_{k+1}^i \approx \gamma_k^i = \frac{D_k^i \cdot R_k^i}{C_k^{i\,2}} \qquad (14)$$

where $\hat{\gamma}_{k+1}^i$ is the estimated value of γ_{k+1}^i.

Combining Eq.11, Eq.13 and Eq.14 together, the allocation bitrate can be derived as,

$$\hat{R}_{k+1}^i = \frac{\dfrac{\gamma_{k+1}^i \cdot C_{k+1}^{i\,2}}{\hat{D}_{k+1}}}{\sum_{j=1}^N \dfrac{\gamma_{k+1}^j \cdot C_{k+1}^{j\,2}}{\hat{D}_{k+1}}} \cdot R_c \qquad (15)$$

which can be rewritten as,

$$\hat{R}_{k+1}^i = \frac{X_{k+1}^i}{\sum_{j=1}^N X_{k+1}^j} \cdot R_c$$

where,

$$X_{k+1}^i = \frac{D_k^i \cdot R_k^i \cdot C_{k+1}^{i\,2}}{C_k^{i\,2}} \qquad (16)$$

Experimental results show that RDAS-I outperforms CAS in [12]. The advantage by applying the inverse proportion model is that we can remove $\hat{D}_{k+1}$ in Eq.15. Otherwise, $\hat{D}_{k+1}$ must be estimated to calculate $\hat{R}_{k+1}^i$. However, the inverse proportion based R-D model is not precious enough to describe the relationship between R and D, with which the allocation error may be amplified when scene changes.

Considering the side effect when applying the inverse proportion model, a more precise R-D model is preferred to conduct a better allocation scheme. In this paper, we propose a refined RDAS in which the R-D relationship is modeled by the Hyperbolic function which outperforms other models on HEVC for most cases [22][23].

The Hyperbolic function based R-D model is one of the highly effective R-D models to represent the characteristics of sequences [20][21] as,

$$D = \alpha \cdot R^\beta \qquad (17)$$

where α and β are parameters related to video contents. Considering R-D model in Eq.17, we use the Hyperbolic function based R-D model for each sequence in a super GOP, then we can have the R-D relationship for each sequence in the $(k+1)_{th}$ super GOP as,

$$\hat{R}_{k+1}^i = \alpha_{k+1}^i \cdot \hat{D}_{k+1}^{\beta_{k+1}^i} \qquad (18)$$

where α_{k+1}^i and β_{k+1}^i will be discussed in part D.

Combining Eq.11 and Eq.18 together, the allocation bitrate can be derived as,

$$\hat{R}_{k+1}^i = \frac{\alpha_{k+1}^i \cdot \hat{D}_{k+1}^{\beta_{k+1}^i}}{\sum_{i=1}^N \alpha_{k+1}^i \cdot \hat{D}_{k+1}^{\beta_{k+1}^i}} \cdot R_c \qquad (19)$$

In Eq.19, we find that we need to estimate α_{k+1}^i, β_{k+1}^i and $\hat{D}_{k+1}$ to obtain $\hat{R}_{k+1}^i$ with a given R_c. Therefore, we will discuss the estimation of $\hat{D}_{k+1}$ in part C and that of $(\alpha_{k+1}^i, \beta_{k+1}^i)$ in part D.

C. A Closed-Form Solution for $\hat{D}_{k+1}$ with the Proposed Joint R-D Model

In Eq.19, one problem caused by applying Hyperbolic model is that $\hat{D}_{k+1}$ must be estimated. A simple way is to brute-force all possible distortions,

$$\hat{D}_{k+1} = D^*, \quad \min_{\{D^* \in [0, MAX^2]\}} \left| R_c - \frac{1}{N} \sum_{i=1}^N \alpha_{k+1}^i \cdot D^{*\beta_{k+1}^i} \right| \qquad (20)$$

where MAX is the maximum possible pixel value of the image. We can accelerate Eq.20 with the binary search algorithm or other algorithms. However, a closed-form solution is more preferred.

To derive a closed-form solution for $\hat{D}_{k+1}$, we define a joint R-D model $\mathcal{R}(D)$ to describe the relationship between the average rate and distortion for a joint stream of multiple sequences. Now, we discuss the joint R-D relationship in one super GOP with the constraint of the ideal allocation in Eq.9,

$$\mathcal{R}(D) = \frac{1}{N} \sum_{i=1}^N \mathcal{R}_i(D) \qquad (21)$$

where the subscript i denotes the i_{th} sequence. D is measured by MSE and $\mathcal{R}(D)$ represents the average BPP of all sequences in one super GOP. Since the MSE of different sequences is the same in the mentioned ideal allocation, $\mathcal{R}(D)$ and $\mathcal{R}_i(D)$ have the same input D.

To derive the joint model $\mathcal{R}(D)$, we analyze the difference of the R-D model between one single sequence and a joint stream. Laplace distribution is commonly used for modeling the distribution of transformed coefficients of natural images [28]. Therefore, we assume that one single transformed coefficient is a zero mean Laplace source $X \sim Laplace(0, b)$ and the probability density of X is,

$$p(x) = \frac{1}{2b} e^{-\frac{|x|}{b}} \qquad (22)$$

According to the information theory [25], we can always find one information channel which can minimize the average mutual information value with the constraint $\bar{D} \leq D$ where $\bar{D}$ is the average distortion of the communication system and D is a constant value. Here we use squared-error distortion. We denote the minimum average interactive information with the constraint $\bar{D} \leq D$ as $R^I(D)$, which can be expressed as

$$R^I(D) = \min_{E(X-\hat{X})^2 \leq D} I(X; \hat{X}) \qquad (23)$$

where $\hat{X}$ is the output value of the channel. $E(X - \hat{X})^2$ is the average distortion $\bar{D}$. $I(X; \hat{X})$ is the average interactive information. $R^I(D)$ is considered as the lower bound of the rate to encode source X with the constraint $\bar{D} \leq D$. According to work [25], we have,

$$I(X; \hat{X}) = h(X) - h(X|\hat{X}) = h(X) - h(X - \hat{X}|\hat{X}) \quad (24)$$
$$\geq h(X) - h(X - \hat{X}) \qquad (25)$$
$$\geq h(X) - h\left(\mathcal{N}\left(0, E(X - \hat{X})^2\right)\right) \qquad (26)$$
$$= h(X) - \frac{1}{2} \ln 2\pi e \bar{D} \qquad (27)$$

where $h(X)$ is the entropy of source X. $h(X|\hat{X})$ is the conditional entropy. $\mathcal{N}\left(0, E(X - \hat{X})^2\right)$ is a zero mean normal distribution with variance $E(X - \hat{X})^2$. Here Eq.25 follows from the fact that conditioning reduces entropy (Theorem 2.6.5 in [25]) and Eq.26 follows from the fact that the normal distribution maximizes the entropy for a given second moment (Theorem 8.6.5 in [25]). According to the definition of entropy, we have,

$$h(X) = -\sum p(x) \ln p(x) = \ln(2be) \qquad (28)$$

where $p(x)$ is the probability density function in Eq.22.

Combine Eq.24-28, we can rewrite Eq.23 as,

$$R^I(D) = \ln(2be) - \frac{1}{2} \ln 2\pi e \bar{D} = \frac{1}{2} \ln\left(\frac{2e}{\pi \bar{D}} \cdot b^2\right) \quad (29)$$

Next, we discuss the calculation of the minimum average interactive information of all transformed coefficients of one stream. In the current video coding system, intra and inter prediction modes are used to eliminate the redundancy among pixels before transforming, and we can consider the transformed coefficients as independent values. Besides, these coefficients may have different distribution parameter b in Eq.22. For example, the distributions of pixels of intra blocks and inter blocks are different. Thus, the problem can be considered as calculating the minimum average interactive information of a parallel Laplace source $X_j \sim Laplace(0, b_j), j = 1,2, \dots, N_P$ where N_P is the number of coefficients in one super GOP. Here b_j might be equal for coefficients with the same mode. The relationship between R^I and D of a parallel Laplace source X_j is,

$$R^I(D) = \frac{1}{2} \sum_{j=1}^{N_P} \ln\left(\frac{2e}{\pi \bar{D}} \cdot b_j^2\right) = \frac{N_P}{2} \ln \frac{2e}{\pi \bar{D}} \cdot \left(\prod_{j=1}^{N_P} b_j\right)^{\frac{2}{N_P}} \qquad (30)$$

According to Eq.30, the R^I-D relationship for coefficients of the same sequence in one super GOP is equivalent to that of a parallel Laplace source $X_j \sim Laplace(0, b'), j = 1,2, \dots, N_P$ where $b' = \left(\prod_{j=1}^{N_P} b_j\right)^{\frac{1}{N_P}}$.

Then, similar with the transformed coefficients of one stream, transformed coefficients of a joint stream in one super GOP can also be considered as a parallel Laplace source $X_{i,j} \sim Laplace(0, b_{i,j}), i = 1,2, \dots, N, j = 1,2, \dots, N_P$ and (i, j) denote the j_{th} coefficient in the i_{th} sequence. Therefore, the $R^I(D)$ of a joint stream $X_{i,j}$ in one super GOP is,

$$R^I(D) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N_P} \ln\left(\frac{2e}{\pi \overline{D}} \cdot b_{i,j}^2\right)$$

$$= \frac{N \cdot N_P}{2} \ln \frac{2e}{\pi \overline{D}} \cdot \left(\prod_{i=1}^{N} \prod_{j=1}^{N_P} b_{i,j}\right)^{\frac{2}{N \cdot N_P}} \qquad (31)$$

Similarly, from Eq.31, the R^I-D relationship for coefficients of the joint stream in one super GOP is consistent with that of a parallel Laplace source $X_{j'} \sim \mathcal{N}(0, b''), j' = 1,2, \dots, N \cdot N_P$ where $b'' = \left(\prod_{i=1}^{N} \prod_{j=1}^{N_P} b_{i,j}\right)^{\frac{1}{N \cdot N_P}}$, that is the R-D relationship of a joint stream can be equivalent to that of one single sequence. The Hyperbolic R-D model in [20] is derived with three assumptions which are satisfied by most natural sequences. If the joint stream still satisfies these assumptions, we can also model the joint R-D relationship with the Hyperbolic function.

To verify these assumptions, we define $h[x]$ as the normalized discrete histogram of the N_P transformed coefficients $a[j], j = 1,2, \dots, N_P$ of one sequence in a super GOP. x is the value of transformed coefficients so that $\sum_x h[x] = 1$. Here N_P is considered to be sufficiently large and the histogram is sufficiently regular. The values of this histogram are interpolated to define a function $p(x) \geq 0$ for all $x \in R$ such that $\int_{-\infty}^{+\infty} p(x) \, dx = 1$. This $p(x)$ is the probability density of a random variable X.

The first assumption is that $p(x)$ is symmetric,

$$p(x) = p(-x) \qquad (32)$$

We sort the N_P transformed coefficients $a[j]$ by their amplitudes. The amplitude of the k_{th} coefficient is written as $a[j_k]$ and $|a[j_k]| \geq |a[j_{k+1}]|, k = 1,2, \dots, N_P$. We define,
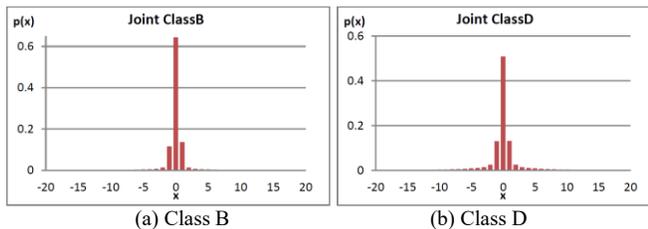

(a) Class B   (b) Class D

Fig.2. The x-$p(x)$ distributions of joint stream for Class B and Class D.
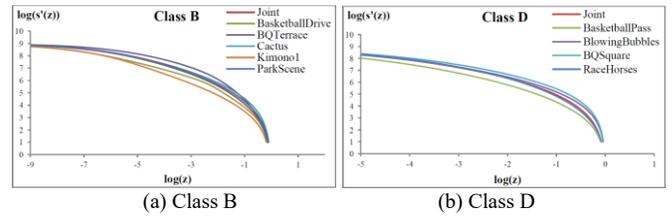

(a) Class B   (b) Class D

Fig.3. The $log_2 z$-$log_2 s'(z)$ curves for Class B and Class D.

$$s\left(\frac{k}{N_P}\right) = |a[j_k]|, \quad k = 1,2, \dots, N_P \qquad (33)$$

Since $0 < \frac{k}{N_P} \leq 1$, we define a function $s'(z)$ for any $z \in [0,1]$ interpolated by the values $s\left(\frac{k}{N_P}\right)$.

The next two assumptions are,

$$\frac{d \, log_2 \, s'(z)}{d \, log_2 \, z} < 0, \qquad z \in [0,1] \qquad (34)$$

$$\frac{d^2 \, log_2 \, s'(z)}{(d \, log_2 \, z)^2} \leq 0, \qquad z \in (0,1) \qquad (35)$$

The three assumptions can be verified by the following analysis.

1) **Assumption 1.** In Fig.2, we give the x-p(x) distributions for the joint streams of Class B and Class D in one super GOP. The distributions are obtained under LowDelay configuration with QP 27 and the coefficients in skipped blocks are all considered as zeros. Moreover, since IDCT is applied in HEVC, coefficients of different TU sizes have been multiplied by different constant values. So, we divide coefficients of TU size 32×32, 16×16, 8×8, and 4×4 by 4, 8, 16, and 32, respectively. The symmetry of the probability density assumption in Eq.32 can be clearly observed and therefore it is verified.

2) **Assumption 2.** In Fig.3, the $log_2 z$-$log_2 s'(z)$ curve for all the sequences of Class B and Class D and their joint streams are shown. As a result, the property of the joint stream curve is similar with that of normal sequences. We can find that $log_2 s'(z)$ strictly decreases when $log_2 z$ increases. So the assumption in Eq.34 is satisfied.

3) **Assumption 3.** Moreover, the curve is obviously concave which means Eq.35 is satisfied.

(a) Class B - LD   (b) Class B - RA
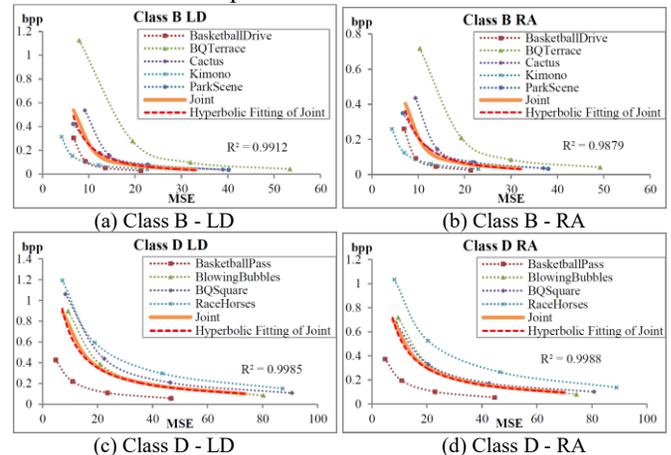
(c) Class D - LD   (d) Class D - RA

Fig.4. The R-D curves for the joint stream and the corresponding hyperbolic fitting curve of Class B and Class D under both LowDelay and RandomAccess configurations.

So, according to the above analysis, the conclusion is that a joint stream of coded streams can also be modeled using a Hyperbolic R-D model and $\mathcal{R}(D)$ in Eq.21 can be modeled as,

$$\mathcal{R}(D) = \frac{1}{N}\sum_{i=1}^{N}\mathcal{R}_i(D) = \alpha \cdot D^{\beta} \qquad (36)$$

We verify our joint R-D model in Eq.36 with some experiments. In Fig.4, we give the R-D curves of one super GOP of Class B and Class D under both LowDelay and RandomAccess configurations. The length of super GOP is set as 16. The horizontal axis and vertical axis denote MSE and BPP, respectively. The dot line is the R-D curve of single sequences and the solid line is the curve of the corresponding joint stream. Besides, we also give the fitting curve of joint stream using Hyperbolic function with dash line and the coefficient of determination R^2 is also given in the figure. As a result, the dash line well fits the solid line which means our proposed Hyperbolic joint model can model the joint R-D relationship well.

Thus, for the k_{th} super GOP, the parameter of joint R-D relationship can be calculated as,

$$\alpha_k \cdot D_k{}^{\beta_k} = \frac{1}{N}\sum_{i=1}^{N}\left(\alpha_k^i \cdot D_k{}^{\beta_k^i}\right) \qquad (37)$$

where α_k and β_k are the parameters related to individual video's content in one super GOP. Considering practical applications, we propose a fast algorithm to calculate α_k and β_k within O(1) time as,

$$\begin{cases} \dfrac{1}{N}\sum_{i=1}^{N}\left(\alpha_k^i \cdot d_k{}^{\beta_k^i}\right) = \hat{\alpha}_k \cdot d_k{}^{\hat{\beta}_k} \\[4mm] \dfrac{1}{N}\sum_{i=1}^{N}\left(\alpha_k^i \cdot (2d_k)^{\beta_k^i}\right) = \hat{\alpha}_k \cdot (2d_k)^{\hat{\beta}_k} \end{cases} \qquad (38)$$

where $d_k = \frac{2}{3}\cdot\overline{D}_k$, and $\overline{D}_k$ is the average value of D_k. $\hat{\alpha}_k$ and $\hat{\beta}_k$ are the estimated values of α_k and β_k. Eq.37 can be rewritten as,

$$\begin{cases} \hat{\beta}_k = \log_2 \dfrac{\sum_{i=1}^{N}\left(\alpha_k^i \cdot (2d_k)^{\beta_k^i}\right)}{\sum_{i=1}^{N}\left(\alpha_k^i \cdot d_k{}^{\beta_k^i}\right)} \\[5mm] \hat{\alpha}_k = \dfrac{1}{N\cdot d_k{}^{\hat{\beta}_k}}\cdot\sum_{i=1}^{N}\left(\alpha_k^i \cdot d_k{}^{\beta_k^i}\right) \end{cases} \qquad (39)$$

To further verify the error of the proposed joint R-D relationship $R_k(\widehat{D}_k) = \hat{\alpha}_k \cdot \widehat{D}_k{}^{\hat{\beta}_k}$ with real R-D values, we define the error function as,

$$\mathcal{E}(D_{Real}) = \frac{\left|N\cdot\hat{\alpha}\cdot D_{Real}{}^{\hat{\beta}} - \sum_{i=1}^{N}R_{Real}^i\right|}{\sum_{i=1}^{N}R_{Real}^i} \qquad (40)$$

where D_{Real} is the real distortion value and R_{Real}^i is the real rate value of the i_{th} stream. $\hat{\alpha}$ and $\hat{\beta}$ are calculated by Eq.39, in which α_k^i and β_k^i are estimated by Least Mean Square (LMS) method with R-D values calculated when QPs are 22, 27, 32, and 37. $\overline{D}_k$ is the average distortion of all sequences obtained with QP 22, 27, 32, and 37.

Experimental results of $\mathcal{E}(D_{Real})$ are calculated on sequences recommended by JCT-VC and the results are given in Table I. LowDelay configuration is applied in HM 16.0. 16 frames are tested with QP 22, 27, 32, and 37. The R-D relationship of each sequence is fitted separately with Hyperbolic function and the corresponding parameters are given in 'α' and 'β' column. '$R(\cdot)$' column denotes the bitrate corresponding to the given distortion $(\cdot)$. $\overline{D}$ is the average MSE of each Class coded with QP 22, 27, 32, and 37. $\overline{D}$ equals to 18, 30, 33, 35, 9, and 11 for Class A to Class F. Line 'Real R of Joint Stream' gives the sum of the real BPP for each class. Line 'Estimate R of Joint Stream' gives the BPP estimated with the proposed joint R-D relationship. Line '$\mathcal{E}(D_{Real})$' gives the error defined in Eq.40. Note that the errors of $R\left(\frac{2}{3}\cdot\overline{D}\right)$ and $R\left(\frac{4}{3}\cdot\overline{D}\right)$ always equal to 0 because of Eq.38. As shown in the experimental results, the error $\mathcal{E}(D_{Real})$ can almost be neglected.

Since we have derived the joint R-D relationship, considering that the channel should be fully utilized for each super GOP, we have,

$$R_c = R_k(\widehat{D}_k) = \hat{\alpha}_k \cdot \widehat{D}_k{}^{\hat{\beta}_k} \qquad (41)$$

So, we can calculate $\widehat{D}_{k+1}$ as,

$$\widehat{D}_{k+1} = \left(\frac{R_c}{\hat{\alpha}_{k+1}}\right)^{\frac{1}{\hat{\beta}_{k+1}}} \qquad (42)$$

$\widehat{R}_{k+1}^i$ in Eq.19 is finally derived as,

$$\widehat{R}_{k+1}^i = R_c \cdot \frac{\alpha_{k+1}^i \cdot \left(\frac{R_c}{\hat{\alpha}_{k+1}}\right)^{\frac{\beta_{k+1}^i}{\hat{\beta}_{k+1}}}}{\sum_{i=1}^{N}\left[\alpha_{k+1}^i \cdot \left(\frac{R_c}{\hat{\alpha}_{k+1}}\right)^{\frac{\beta_{k+1}^i}{\hat{\beta}_{k+1}}}\right]} \qquad (43)$$

In our proposed RDAS-H, bitrate is allocated according to Eq.43. Besides, our model can be directly used under different coding structures or different rate control schemes such as [23] or [33].

D. Parameter Estimation for RDAS-H

Since there is no closed-form expression derived for α and β in the Hyperbolic model, we need to estimate α_{k+1}^i and β_{k+1}^i by the obtained value of the k_{th} super GOP to allocate bitrate for the $(k+1)_{th}$ super GOP in Eq.43.

TABLE I THE VERIFICATION OF THE ESTIMATION IN Eq.20 OF SEQUENCES RECOMMENDED BY JCT-VC.

Class	Sequence	α	β	$R(\frac{1}{2}\cdot\bar{D})$	$R(\frac{2}{3}\cdot\bar{D})$	$R(\frac{5}{6}\cdot\bar{D})$	$R(\bar{D})$	$R(\frac{7}{6}\cdot\bar{D})$	$R(\frac{4}{3}\cdot\bar{D})$	$R(\frac{3}{2}\cdot\bar{D})$
ClassA	PeopleOnStreet	1.688	-0.944	0.21216	0.16171	0.13099	0.11028	0.09535	0.08405	0.07521
	Traffic	1.044	-1.250	0.06699	0.04675	0.03537	0.02816	0.02323	0.01966	0.01697
	Real R of Joint Stream			0.27915	0.20846	0.16636	0.13844	0.11857	0.10371	0.09217
	Estimated R of Joint Stream	1.274	-1.007	0.27852	0.20846	0.16650	0.13857	0.11864	0.10371	0.09211
	$\mathcal{E}(D_{Real})$			**0.22%**	**0.00%**	**0.08%**	**0.09%**	**0.06%**	**0.00%**	**0.07%**
ClassB	BasketballDrive	0.887	-0.998	0.05951	0.04466	0.03574	0.02980	0.02555	0.02236	0.01988
	BQTerrace	26.822	-1.794	0.20825	0.12429	0.08329	0.06005	0.04554	0.03584	0.02902
	Kimono1	4.469	-0.975	0.31877	0.24080	0.19372	0.16217	0.13954	0.12250	0.10921
	ParkScene	3.339	-0.832	0.35083	0.27615	0.22936	0.19708	0.17336	0.15513	0.14065
	Cactus	13.165	-1.829	0.09297	0.05493	0.03652	0.02617	0.01974	0.01546	0.01247
	Real R of Joint Stream			1.03032	0.74084	0.57864	0.47527	0.40373	0.35130	0.31123
	Estimated R of Joint Stream	3.726	-1.076	1.00975	0.74084	0.58264	0.47882	0.40560	0.35130	0.30947
	$\mathcal{E}(D_{Real})$			**2.00%**	**0.00%**	**0.69%**	**0.75%**	**0.46%**	**0.00%**	**0.56%**
ClassC	BasketballDrill	1.382	-1.084	0.06620	0.04846	0.03805	0.03123	0.02642	0.02286	0.02012
	BQMall	1.562	-1.007	0.09282	0.06948	0.05549	0.04619	0.03955	0.03457	0.03070
	PartyScene	5.685	-0.890	0.46900	0.36306	0.29767	0.25308	0.22064	0.19591	0.17642
	RaceHorses	4.885	-1.020	0.27990	0.20872	0.16623	0.13802	0.11794	0.10292	0.09127
	Real R of Joint Stream			0.90793	0.68972	0.55745	0.46852	0.40455	0.35627	0.31851
	Estimated R of Joint Stream	3.281	-0.953	0.90729	0.68972	0.55759	0.46865	0.40462	0.35627	0.31844
	$\mathcal{E}(D_{Real})$			**0.07%**	**0.00%**	**0.03%**	**0.03%**	**0.02%**	**0.00%**	**0.02%**
ClassD	BasketballPass	0.887	-0.998	0.07138	0.05357	0.04288	0.03574	0.03065	0.02683	0.02385
	BlowingBubbles	5.647	-1.110	0.34216	0.24862	0.19408	0.15852	0.13359	0.11519	0.10107
	BQSquare	4.469	-0.975	0.38078	0.28765	0.23140	0.19372	0.16669	0.14634	0.13046
	RaceHorses	3.339	-0.832	0.40830	0.32138	0.26693	0.22936	0.20175	0.18054	0.16369
	Real R of Joint Stream			1.20261	0.91122	0.73529	0.61734	0.53267	0.46889	0.41907
	Estimated R of Joint Stream	1.784	-0.792	1.20057	0.91122	0.73575	0.61777	0.53291	0.46889	0.41883
	$\mathcal{E}(D_{Real})$			**0.17%**	**0.00%**	**0.06%**	**0.07%**	**0.04%**	**0.00%**	**0.06%**
ClassE	FourPeople	0.23	-1.035	0.04771	0.03542	0.02812	0.02328	0.01985	0.01729	0.01530
	Johnny	0.43	-1.691	0.03392	0.02086	0.01430	0.01051	0.00810	0.00646	0.00529
	KristenAndSara	0.22	-1.246	0.03346	0.02338	0.01771	0.01411	0.01164	0.00986	0.00851
	Real R of Joint Stream			0.11510	0.07966	0.06013	0.04790	0.03959	0.03360	0.02911
	Estimated R of Joint Stream	0.143	-1.030	0.11398	0.07966	0.06034	0.04808	0.03968	0.03360	0.02902
	$\mathcal{E}(D_{Real})$			**0.97%**	**0.00%**	**0.35%**	**0.38%**	**0.24%**	**0.00%**	**0.30%**
ClassF	ChinaSpeed	0.494	-0.666	0.08132	0.06714	0.05787	0.05125	0.04625	0.04231	0.03912
	SlideEditing	0.103	-0.265	0.05040	0.04670	0.04402	0.04194	0.04026	0.03886	0.03767
	SlideShow	0.063	-0.593	0.01267	0.01068	0.00936	0.00840	0.00766	0.00708	0.00660
	Real R of Joint Stream			0.14438	0.12452	0.11124	0.10159	0.09418	0.08826	0.08339
	Estimated R of Joint Stream	0.159	-0.440	0.14364	0.12452	0.11146	0.10181	0.09431	0.08826	0.08325
	$\mathcal{E}(D_{Real})$			**0.52%**	**0.00%**	**0.19%**	**0.22%**	**0.14%**	**0.00%**	**0.18%**

As we have mentioned before, the rate control algorithm in [23] is applied in our system, in which a Hyperbolic model is also used. Thus, the result of the Rate Controller can be reused. In [23], the method to estimate parameters of the Hyperbolic model is proposed, while the estimation method is further improved in the recent work [27] and we use the parameter estimation method of [27] instead of this part of [23] in our Rate Controller.

The Hyperbolic model used in [23][27] is,

$$\lambda = A \cdot R^B \tag{44}$$

where A and B are parameters related to video contents. And,

$$\lambda = -\frac{dD}{dR} \qquad (45)$$

After the encoding of the k_{th} super GOP, we can obtain A and B from the Rate Controller. Combining Eq.44-45 and the model in Eq.17, we can have the following relationship of α, β and A, B as,

$$\begin{cases} \alpha_{k,p}^i = \left(-A_{k,p}^i \cdot \beta_{k,p}^i\right)^{-\beta_{k,p}^i} \\ \beta_{k,p}^i = \dfrac{1}{B_{k,p}^i + 1} \end{cases} \qquad (46)$$

where i is the sequence index, k is the index of the super GOP, and p denotes the p_{th} frame in one super GOP. In case that Hyperbolic model is not used in the Rate Controller and we cannot obtain $A_{k,p}^i$ and $B_{k,p}^i$, we give another solution to calculate $\alpha_{k,p}^i, \beta_{k,p}^i$ with $R_{k,p}^i, D_{k,p}^i$ by a similar algorithm in [27]. Considering the definition of the Lagrange multiplier and the Hyperbolic R-D relationship, we have two equations,

$$\begin{cases} \lambda_{k,p}^i = -\dfrac{dD_{k,p}^i}{dR_{k,p}^i} = \dfrac{1}{-\alpha_{k,p}^i \beta_{k,p}^i \cdot D_{k,p}^{i\,\beta_{k,p}^i - 1}} \\ R_{k,p}^i = \alpha_{k,p}^i \cdot D_{k,p}^{i\,\beta_{k,p}^i} \end{cases} \qquad (47)$$

Since λ_k^i, $R_{k,p}^i$ and $D_{k,p}^i$ can always be obtained after the frame is encoded, we calculate $\alpha_{k,p}^i$ and $\beta_{k,p}^i$ as,

$$\begin{cases} \beta_{k,p}^i = -\dfrac{D_{k,p}^i}{\lambda_{k,p}^i \cdot R_{k,p}^i} \\ \alpha_{k,p}^i = \dfrac{R_{k,p}^i}{D_{k,p}^{i\,\beta_{k,p}^i}} \end{cases} \qquad (48)$$

As given in Eq.47-48 or Eq.46, we can obtain $\alpha_{k,p}^i$ and $\beta_{k,p}^i$ for each frame. However, how to calculate α_k^i and β_k^i as the parameters of the R-D relationship for all the frames in one super GOP still remains unsolved. We cannot calculate α_k^i and β_k^i by the average value of the whole super GOP because λ is different for each frame. Therefore, we calculate α_k^i and β_k^i with a similar way as that we have mentioned in Eq.39,

$$\begin{cases} \hat{\beta}_k^i = \log_2 \dfrac{\frac{1}{M}\sum_{p=1}^M \left(\alpha_{k,p}^i \cdot \left(2D_{k,p}^i\right)^{\beta_{k,p}^i}\right)}{\bar{R}_k^i} \\ \hat{\alpha}_k^i = \dfrac{\bar{R}_k^i}{M \cdot \bar{D}_k^{i\,\hat{\beta}_k}} \end{cases} \qquad (49)$$

where M is the number of frames in a super GOP. $\bar{R}_k^i$, $\bar{D}_k^i$ are the average BPP and MSE for the whole super GOP.

To further estimate α_{k+1}^i and β_{k+1}^i, the only information we can obtain is the complexity measure in Eq.5-7 calculated with

look-ahead approach. Therefore, we try to detect the relationship of α, β and the complexity measure C in Eq.5 by some experimental results. We test all sequences from Class A to F under both RandomAccess and LowDelay configurations and the results are all shown in Fig.5. The horizontal axis denotes C and vertical axis denotes α or β. Each point represents C, α and β of one super GOP. The length of one super GOP is set as 16 and each intra-frame is inserted as the first frame of every super GOP. We calculated $\hat{\alpha}_k^i$ and $\hat{\beta}_k^i$ by Eq.49 and complexity measure C by Eq.7 for each super GOP. Since the parameter estimation is not our main work in this paper, we use a simple and efficient linear function to express the relationship between α and C according to the results shown in Fig.5. What's more, the relationship of β and C looks independent to some extent. Therefore, we estimate α_{k+1}^i and β_{k+1}^i by,

$$\begin{cases} \hat{\alpha}_{k+1}^i = \dfrac{\bar{C}_{k+1}^i}{\bar{C}_k^i} \hat{\alpha}_k^i \\ \hat{\beta}_{k+1}^i = \hat{\beta}_k^i \end{cases} \qquad (50)$$

where $\hat{\alpha}_{k+1}^i$ and $\hat{\beta}_{k+1}^i$ are the estimated value of α_{k+1}^i and β_{k+1}^i. According to our experimental results in the next Section, we find this simple estimation performs better than directly estimating $\alpha_{k+1}^i, \beta_{k+1}^i$ with $\hat{\alpha}_k^i, \hat{\beta}_k^i$.
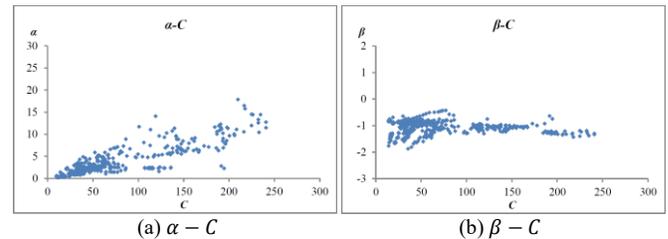
Fig.5. The relationship between the parameter of Hyperbolic R-D model and the complexity measure.

E. Rate-Distortion Performance Analysis for RDAS-H

Although the proposed scheme focuses on minimizing the variance among multiple sequences, we are still interested in the R-D performance of RDAS-H.

To analyze the R-D performance, we assume the parameters for the Hyperbolic R-D model are accurate and the model can precisely express the R-D relationship. To compare with RDAS-H, we allocate bitrate for different sequences equally and set this allocation method as the anchor named AVG. Then we have,

$$\begin{cases} R_C = \alpha_i \cdot D_i^{\beta_i} & (51) \\ R_c = \alpha \cdot D^{\beta} & (52) \end{cases}$$

where Eq.51 and Eq.52 describe the R-D relationship of applying AVG and RDAS-H as allocation scheme, respectively. In Eq.51, the subscript i denotes the sequence index and α_i, β_i are the parameters of each sequence. Eq.51 means that each sequence has a different distortion but the same bitrate. In

Eq.52, α, β are the parameters for a joint R-D relationship. Eq.52 means that each sequence has the same distortion and a different bitrate, while the overall bitrate is R_c.

Since the overall bitrate of two schemes are the same, we only need to compare the distortion,

$$\begin{cases} D - \frac{1}{N}\sum_{i=1}^{N} D_i \leq 0, & RDAS-H \ is \ better \\ D - \frac{1}{N}\sum_{i=1}^{N} D_i \geq 0, & AVG \ is \ better \end{cases} \qquad (53)$$

So, we need to find out whether $D - \frac{1}{N}\sum_{i=1}^{N} D_i$ is positive or negative. $D - \frac{1}{N}\sum_{i=1}^{N} D_i$ can also be expressed as,

$$D - \frac{1}{N}\sum_{i=1}^{N} D_i = \left(\frac{R_c}{\alpha}\right)^{\frac{1}{\beta}} - \frac{1}{N}\sum_{i=1}^{N}\left(\frac{R_c}{\alpha_i}\right)^{\frac{1}{\beta_i}} \qquad (54)$$

According to Eq.39, α, β can be calculated with α_i and β_i, so the R-D performance will totally depend on α_i, β_i and R_c. Let us take the parameters of the hyperbolic R-D model for different sequences and the joint stream in Table I as an example to calculate $D - \frac{1}{N}\sum_{i=1}^{N} D_i$, and we give the results in Table II. Different R_c values are tested and column 'R' denotes the bitrate $R(\overline{D})$ given in Table I. We find that in most situations $D - \frac{1}{N}\sum_{i=1}^{N} D_i$ is negative, which means RDAS-H is better than AVG. Since Eq.51 is not always negative, the R-D performance is affected by the characteristics of the sequences and the overall bandwidth. However, as a result shown in Table II, the R-D performance of our proposed RDAS-H will be better than AVG in most cases (the usually usable range), and we can also draw a conclusion that our RDAS-H will not have a bad R-D performance compared to AVG in most cases (and also the typical situations).

TABLE II THE THEORETICAL R-D PERFORMANCE OF SEQUENCES RECOMMENDED BY JCT-VC.

Sequence	1/4R	1/3R	1/2R	R	2R	3R	4R
ClassA	-5.86	-4.06	-2.45	-1.09	-0.53	-0.37	-0.29
ClassB	-33.51	-21.66	-11.61	-4.03	-1.6	-1.08	-0.89
ClassC	-11.14	-7.74	-4.65	-1.99	-0.89	-0.57	-0.42
ClassD	11.65	3.93	-1.51	-3.82	-3.2	-2.58	-2.15
ClassE	1.89	0.64	-0.4	-1.06	-1.08	-0.98	-0.89
ClassF	-8.57	-2.79	-1.55	-1.11	-0.57	-0.35	-0.24

IV. EXPERIMENTAL RESULTS

We integrate the proposed RDAS-H into HEVC reference software HM 16.0. Both LowDelay and RandomAccess configurations are used. The λ-domain rate control algorithm in [23] is applied and only frame-level rate control is turned on for all the methods in this paper (including the proposed scheme and related schemes). Test sequences recommended by JCT-VC containing 6 classes are tested. The length of a super

GOP is set as 16. In our experiments, we jointly encode all sequences in one class each time and the sequence numbers for Class A to Class F are 2, 5, 4, 4, 3, and 3. 288 frames are tested for Class C, Class D, Class E, and Class F. 240 frames are tested for Class B and 144 frames for Class A. Intra Period is set as 16. In each experiment, we encode sequences with fixed QPs in advance, and then we calculate R_c by summing up the bitrate of each sequence with the corresponding fixed QPs.

Variance measures how far a set of numbers are spread out from their mean. According to the target function described in Eq.2, a smaller variance of distortion represents a better allocation model. Thus, the evaluation index for the $minVAR$ problem is defined as,

$$Variance_k = \frac{1}{N}\sum_{i=1}^{N}\left(PSNR_k^i - \overline{PSNR_k}\right)^2 \qquad (55)$$

where $Variance_k$ represents the variance of Peak Signal to Noise Ratio (PSNR) of the k_{th} super GOPs of all the sequences. $PSNR_k^i$ represents the PSNR value of the k_{th} super GOP in the i_{th} video and $\overline{PSNR_k}$ is the mean value of PSNRs of different sequences in the k_{th} super GOP.

To verify the performance of our proposed scheme, considering the target of our multiplexing problem is to minimize the variance of distortion, we define the variance saving ratio (VSR) between our proposed scheme and the anchor as,

$$VSR_{anchor} = \frac{Variance_{anchor} - Variance_{proposed}}{Variance_{anchor}} \qquad (56)$$

where $Variance_{anchor}$ and $Variance_{proposed}$ are the average variance values of all super GOPs for the anchor and the proposed scheme, respectively. A larger VSR_{anchor} represents a better performance of the proposed scheme compared with the anchor.

The main objectives of our experiments are two-folds: (1) to explore the effectiveness of our proposed methods described in Part C and Part D in Section III, and (2) to evaluate the performance of our solution (RDAS-H) compared with the related scheme in [12] and that in our previous work [17].

1) Proposed methods verification:

In this set of experiments, the main objective is to verify the estimation accuracy of α_{k+1}^i, β_{k+1}^i and $\widehat{D}_{k+1}$ in Eq.19.

Verification of $\widehat{D}_{k+1}$ estimation. We first evaluate our proposed closed-form solution of $\widehat{D}_{k+1}$, which is derived in Eq.42. We compare our estimation with the brute-force method described in Eq.20 and the brute-force method can achieve a better performance theoretically. These two methods are both integrated into our RDAS-H and the other configurations are set the same. QP value is set as 32, and the comparison results are given in Table III. Our proposed estimation is denoted as 'Ours', while the brute-force method is denoted as 'Brute-force'. We also give the VSR performance in the column 'VSR' where 'Brute-force' is set as the anchor. A

negative VSR indicates a better performance of the brute-force method than our estimation for $\widehat{D}_{k+1}$. Our estimation has a negligible performance loss compared with the theoretical best performance, which proves the efficiency of our method.

Verification of α_{k+1}^i and β_{k+1}^i. Then, we verify the parameters estimation for the Hyperbolic model. In Part D Section III, we derive the relationship between parameter (α, β) with complexity measure C by some experimental results and then estimate α_{k+1}^i and β_{k+1}^i in Eq.50. We also compare our estimation method with a direct estimation method which estimates $(\alpha_{k+1}^i, \beta_{k+1}^i)$ with $(\hat{\alpha}_k^i, \hat{\beta}_k^i)$ in Eq.49. The results are shown in Table IV where our proposed estimation method is denoted as 'Ours' while the direct estimation method is denoted as 'Direct estimation'. The experimental environment is the same as that in Table III and 'Direct estimation' is set as the anchor when calculating VSR. Our estimation in Eq.50 definitely improves the performance by saving 14.60% and 9.75% variance on average under LowDelay and RandomAccess configurations respectively.

2) Entire solution (RDAS-H) verification:

In Table V, we verify the performance of the proposed RDAS-H method compared with three other methods. Four different total bitrates are tested and bitrate for each class is given in column 'Rate'. In column 'Scheme', 'AVG' represents that each stream is allocated with the same bitrate. 'CAS[12]' is the method given in [12] which is briefly described in Eq.4. 'RDAS-H' is our proposed method and 'RDAS-I[17]' is our previous method in [17] applying the same complexity measure with that in 'CAS[12]'. We apply the rate control algorithm [23] in all these methods. Besides, 'VSR AVG/CAS/RDAS-I' gives the VSR between RDAS-H and AVG/CAS/RDAS-I, respectively. Since the first super GOP is allocated with the same bitrate among all three methods, the variance of the first super GOP is not considered for the performance evaluation.

To analyze the results, we first focus on VSR in line 'VSR AVG' and 'VSR CAS[12]'. Our proposed method significantly outperforms AVG and CAS [12] with VSR 82.63% and 75.29% on average under different bitrates. Moreover, comparing with our previous method RDAS-I [17], the result is still considerable with VSR 36.62%. As the overall bitrate decreasing, VSR for the Hyperbolic increases because the inverse proportion R-D model is not precise enough. Because when the overall bitrate is low, distortion will change more while bitrate changes a little and Hyperbolic R-D model describes this relationship better. Therefore, the more precise Hyperbolic R-D model will be more effective at low bandwidth.

To clearly illustrate the results in Table V, we give the variance curves of AVG, CAS[12], RDAS-I[17], and RDAS-H in Fig.6. The curves show the average variance value of each frame corresponding to the results of the third group in Table V. The horizontal and vertical axes denote the frame number and the variance, respectively. In the title, 'LD' and 'RA' denote the LowDelay and RandomAccess configurations. As shown in Fig.6, the proposed RDAS-H outperforms RDAS-I [17] method and the related CAS [12] method in almost all super

GOPs.

Besides QP values commonly used, our proposed RDAS-H can also work with a lower or higher QP. In Fig.7, the VSR values between RDAS-H and AVG/CAS/RDAS-I are given with QP 12, 17, 22, 27, 32, 37, 42, and 47. According to our results, our scheme can work under different QPs. The efficiency of our proposed scheme relies on the accuracy of R-D model, and the accuracy of R-D model depends on the estimation of parameters (e.g. α, β in the Hyperbolic model) in Rate Controller. Based on the authors experience, α, β are difficult to estimate when QP is very low or very high which may affect our allocation efficiency to some extent. The estimation of α, β is not our main purpose in this paper, and we will focus on this part in our future work.

According to our target function, our scheme only focuses on minimizing the variance among sequences. However, we still want to know how the proposed RDAS-H affects the R-D performance. We evaluate our solution with BD-BR to further prove the effectiveness of our solution. Therefore, we give the BD-BR performance of CAS [12], RDAS-I [17], and RDAS-H in Table VI, where AVG is set as the anchor. The first

TABLE III THE ACCURACY OF $\widehat{D}_{k+1}$ ESTIMATION OF OUR METHOD AND BRUTE-FORCE.

Configuration	Sequence	Variance		VSR
		Brute-force	Ours	
LowDelay	ClassA	1.84	1.82	1.30%
	ClassB	0.42	0.43	-2.87%
	ClassC	0.91	0.91	-0.44%
	ClassD	0.44	0.44	-0.92%
	ClassE	0.19	0.19	-2.15%
	ClassF	17.68	17.68	0.02%
	Average	3.58	3.58	**-0.84%**
RandomAccess	ClassA	4.15	4.12	0.77%
	ClassB	0.65	0.66	-1.85%
	ClassC	2.34	2.34	-0.17%
	ClassD	0.77	0.77	0.00%
	ClassE	0.22	0.23	-3.60%
	ClassF	11.29	11.29	-0.04%
	Average	3.24	3.24	**-0.82%**

TABLE IV THE ACCURACY OF α_{k+1}^i AND β_{k+1}^i ESTIMATION OF OUR METHOD AND DIRECT ESTIMATION.

Configuration	Sequence	Variance		VSR
		Direct estimation	Ours	
LowDelay	ClassA	2.21	1.82	17.46%
	ClassB	0.59	0.43	26.50%
	ClassC	1.07	0.91	14.95%
	ClassD	0.51	0.44	12.87%
	ClassE	0.23	0.19	15.56%
	ClassF	17.73	17.68	0.25%
	Average	3.72	3.58	**14.60%**
RandomAccess	ClassA	5.16	4.12	20.16%
	ClassB	0.72	0.66	8.33%
	ClassC	2.58	2.34	9.30%
	ClassD	0.82	0.77	5.52%
	ClassE	0.27	0.23	14.81%
	ClassF	11.34	11.29	0.40%
	Average	3.48	3.24	**9.75%**

TABLE V
COMPARISONS OF THE VARIANCES OF CAS[12], RDAS-I[17], AND RDAS-H.

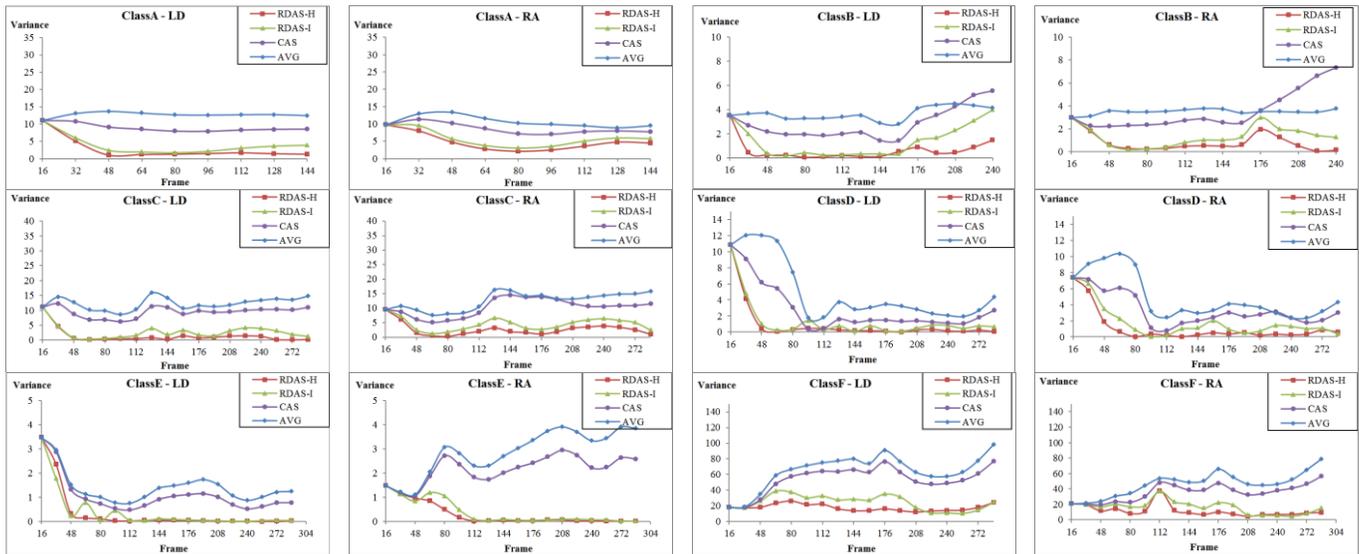| QP	Configuration	Scheme	ClassA	ClassB	ClassC	ClassD	ClassE	ClassF	Average
22	Low Delay	AVG	5.33	4.03	9.53	6.96	0.61	44.34	11.80
		CAS[12]	2.60	2.81	4.99	2.68	0.45	33.58	7.85
		RDAS-I[17]	1.06	0.93	1.87	1.04	0.23	14.55	3.28
		RDAS-H	0.89	0.39	1.07	0.73	0.18	10.03	2.21
		VSR AVG	**83.39%**	**90.36%**	**88.79%**	**89.55%**	**70.98%**	**77.37%**	**83.41%**
		VSR CAS	**65.99%**	**86.16%**	**78.60%**	**72.83%**	**60.88%**	**70.12%**	**72.43%**
		VSR RDAS-I	**16.21%**	**58.45%**	**42.99%**	**30.04%**	**23.14%**	**31.01%**	**33.64%**
	Random Access	AVG	4.80	3.34	7.11	5.49	0.86	30.54	8.69
		CAS[12]	2.57	3.52	3.28	2.75	0.59	20.37	5.51
		RDAS-I[17]	1.61	0.68	0.99	0.95	0.24	8.57	2.17
		RDAS-H	1.28	0.43	0.54	0.51	0.19	7.18	1.69
		VSR AVG	**73.44%**	**87.03%**	**92.38%**	**90.78%**	**77.45%**	**76.50%**	**82.93%**
		VSR CAS	**50.35%**	**87.70%**	**83.50%**	**81.62%**	**67.11%**	**64.77%**	**72.51%**
		VSR RDAS-I	**20.90%**	**36.35%**	**45.34%**	**46.84%**	**19.09%**	**16.31%**	**30.80%**
27	Low Delay	AVG	7.90	3.91	11.42	2.92	0.85	56.42	13.90
		CAS[12]	4.72	3.02	7.23	2.55	0.61	44.69	10.47
		RDAS-I[17]	1.88	1.13	1.99	0.93	0.22	19.37	4.25
		RDAS-H	1.56	0.65	1.13	0.57	0.17	13.81	2.98
		VSR AVG	**80.28%**	**83.26%**	**90.10%**	**80.42%**	**80.35%**	**75.53%**	**81.66%**
		VSR CAS	**67.00%**	**78.31%**	**84.35%**	**77.56%**	**72.61%**	**69.10%**	**74.82%**
		VSR RDAS-I	**17.28%**	**42.01%**	**43.15%**	**38.66%**	**23.45%**	**28.72%**	**32.21%**
	Random Access	AVG	6.74	3.74	10.42	5.23	1.71	38.38	11.04
		CAS[12]	4.22	3.39	6.34	2.92	1.18	27.96	7.67
		RDAS-I[17]	2.86	1.17	1.79	1.17	0.23	12.10	3.22
		RDAS-H	2.27	0.81	1.09	0.74	0.19	9.13	2.37
		VSR AVG	**66.27%**	**78.26%**	**89.51%**	**85.75%**	**88.77%**	**76.22%**	**80.80%**
		VSR CAS	**46.19%**	**76.00%**	**82.74%**	**74.53%**	**83.72%**	**67.34%**	**71.75%**
		VSR RDAS-I	**20.67%**	**30.21%**	**39.05%**	**36.28%**	**17.78%**	**24.54%**	**28.09%**
32	Low Delay	AVG	12.90	3.65	12.35	4.63	1.31	67.15	17.00
		CAS[12]	8.71	2.79	9.40	2.41	0.95	55.78	13.34
		RDAS-I[17]	3.09	1.20	2.29	0.83	0.22	24.88	5.42
		RDAS-H	1.82	0.43	0.91	0.44	0.19	17.68	3.58
		VSR AVG	**85.88%**	**88.29%**	**92.65%**	**90.40%**	**85.46%**	**73.66%**	**86.06%**
		VSR CAS	**79.10%**	**84.67%**	**90.35%**	**81.54%**	**79.99%**	**68.30%**	**80.66%**
		VSR RDAS-I	**41.13%**	**64.37%**	**60.31%**	**46.08%**	**15.10%**	**28.91%**	**42.65%**
	Random Access	AVG	10.73	3.53	12.73	4.71	2.94	47.82	13.74
		CAS[12]	8.49	3.56	10.37	3.11	2.21	36.47	10.70
		RDAS-I[17]	5.28	1.21	4.27	1.46	0.32	16.15	4.78
		RDAS-H	4.12	0.66	2.34	0.77	0.23	11.29	3.24
		VSR AVG	**61.55%**	**81.26%**	**81.59%**	**83.72%**	**92.01%**	**76.38%**	**79.42%**
		VSR CAS	**51.43%**	**81.46%**	**77.41%**	**75.33%**	**89.38%**	**69.03%**	**74.01%**
		VSR RDAS-I	**21.84%**	**45.34%**	**45.19%**	**47.60%**	**26.46%**	**30.09%**	**36.09%**
37	Low Delay	AVG	24.74	2.56	13.74	3.78	2.30	78.50	20.94
		CAS[12]	17.68	2.23	10.79	1.93	1.72	64.85	16.53
		RDAS-I[17]	4.45	1.32	2.43	1.15	0.27	29.34	6.49
		RDAS-H	1.43	0.62	0.92	0.38	0.20	22.95	4.41
		VSR AVG	**94.23%**	**75.88%**	**93.33%**	**90.05%**	**91.33%**	**70.76%**	**85.93%**
		VSR CAS	**91.92%**	**72.34%**	**91.50%**	**80.54%**	**88.42%**	**64.61%**	**81.56%**
		VSR RDAS-I	**67.94%**	**53.32%**	**62.30%**	**67.24%**	**25.94%**	**21.78%**	**49.75%**
	Random Access	AVG	20.89	3.29	14.61	4.72	4.34	57.57	17.57
		CAS[12]	15.94	2.39	13.97	2.88	3.40	42.72	13.55
		RDAS-I	9.74	1.02	4.81	1.45	0.53	20.57	6.35
		RDAS-H	7.21	0.50	2.58	0.73	0.36	13.64	4.17
		VSR AVG	**65.50%**	**84.69%**	**82.33%**	**84.62%**	**91.64%**	**76.30%**	**80.85%**
		VSR CAS	**54.80%**	**78.98%**	**81.51%**	**74.83%**	**89.31%**	**68.07%**	**74.58%**
		VSR RDAS-I	**26.05%**	**50.66%**	**46.30%**	**49.80%**	**31.90%**	**33.69%**	**39.73%**

:

Fig.6. Comparisons of variance curves of AVG, CAS[12], RDAS-I[17], and RDAS-H for Class A-F under both LowDelay and RandomAccess configurations.

TABLE VI COMPARISONS OF BD-BR OF CAS[12], RDAS-I[17], AND RDAS-H.

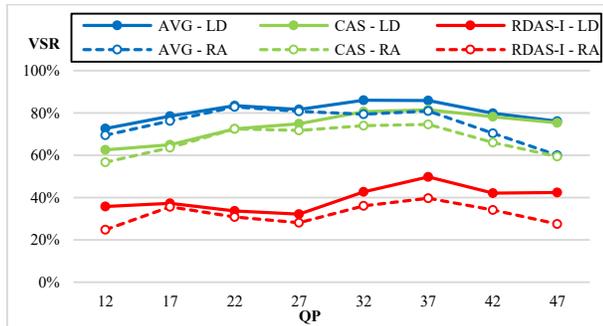| Configuration	Scheme	ClassA	ClassB	ClassC	ClassD	ClassE	ClassF	Average
LowDelay	CAS[12]	-2.23%	-4.34%	-1.71%	1.37%	-0.59%	-1.85%	**-1.56%**
	RDAS-I[17]	-2.17%	-1.78%	-0.28%	0.11%	-0.20%	-2.33%	**-1.11%**
	RDAS-H	-2.29%	-2.32%	-0.94%	0.43%	0.35%	-2.29%	**-1.18%**
RandomAccess	CAS[12]	-3.80%	-0.90%	-2.66%	1.56%	-0.50%	-2.54%	**-1.47%**
	RDAS-I[17]	-4.92%	-2.19%	-2.38%	1.09%	0.63%	-1.22%	**-1.50%**
	RDAS-H	-5.17%	-1.70%	-3.23%	1.40%	0.87%	-1.02%	**-1.48%**

Fig.7. The VSR values with different QPs of the proposed RDAS-H compared with AVG, CAS[12], and RDAS-I[17].

super GOP is excluded from the results because the first super GOP is allocated with the same bitrate for all the methods. The experimental R-D performance of our proposed RDAS-H is better than that of AVG in most cases and the BD-BRs of CAS[12], RDAS-I[17], and our proposed method are comparable. As a result, our method gets a better min-VAR performance without RD performance reduction.

Since the computational complexity of the HEVC encoder can be high, many researchers have devoted their efforts on reducing the computational complexity of the HEVC encoder[30-32]. Therefore, we also analyze the computational complexity of our proposed scheme because we don't want to spend too much complexity on the allocation part. According to our allocation method in Eq.43, the time cost is only related to the number of sequences but the resolutions of the sequences. In table VII, we give the time cost of our allocation scheme and

that of the whole HEVC encoder. The time cost for encoding each sequence is given in the 'Encoder (s)' column and the time cost for our proposed allocation scheme is given in the 'RDAS-H (μs)' column. We can find that the encoder will take more time to process a larger resolution sequence but our allocation scheme doesn't. And the time cost of our proposed scheme is only related to the number of sequences. Besides, the time for the allocation is 10^7 times less than that for encoding.

TABLE VII THE TIME COST OF OUR ALLOCATION METHOD AND THAT OF THE WHOLE HEVC ENCODER.

Sequence	Resolution	Number of Sequences	RDAS-H (μs)	HEVC Encoder(s)
Class A	2560×1600	2	**4.515**	2021.02
Class B	1920×1080	5	**10.016**	1705.76
Class C	832×480	4	**8.859**	427.35
Class D	416×240	4	**8.922**	90.72
Class E	1280×720	3	**6.719**	287.62
Class F	1280×720/1024×768	3	**5.610**	407.67

V. CONCLUSION

In this paper, we proposed a novel joint rate-distortion based allocation scheme derived by an ideal allocation situation to solve the *minVAR* problem. The proposed RDAS can work with different R-D models. Furthermore, a Hyperbolic function based R-D model is applied to the RDAS scheme (RDAS-H) to obtain a better allocation result. We derive a closed-form solution of RDAS-H instead of a brute-force solution by deriving a joint R-D model based on information theory. Compared with the related CAS in [12], the proposed RDAS-H improves the performance by saving variance 75.29% on

average under LowDelay and RandomAccess configurations with QP 22, 27, 32, and 37. Besides, the proposed method saves 36.62% variance over our previous method RDAS-I in [17], in which the R-D model is built by an inverse proportion function. Besides, the proposed RDAS-H can also obtain a higher R-D performance compared with the methods which allocate bitrate for different sequences equally in most cases. In this work we emphasize on reaching an equal visual quality among all sequences. However, it will be a better idea if we encode joint videos according to the users' heterogeneous characters and how to transmit joint videos to meet every user's demand is another difficult problem which is part of our future work.

REFERENCES

[1] Z. He; S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling," IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 10, pp. 840-849, Oct. 2002.

[2] M. Tiwari; T. Groves; P. C. Cosman, "Competitive Equilibrium Bitrate Allocation for Multiple Video Streams," IEEE Transactions on Image Processing, vol. 19, no. 4, pp. 1009-1021, April 2010.

[3] C. Pang; O. C. Au; F. Zou; J. Dai, "Model-based optimal dependent joint bit allocation of H.264/AVC statistical multiplexing," IEEE International Conference on Acoustics, Speech and Signal Processing, 2012, pp. 1357-1360, Mar. 2012.

[4] C. Pang; O. C. Au; J. Dai; F. Zou, "Dependent Joint Bit Allocation for H.264AVC Statistical Multiplexing Using Convex Relaxation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 8, pp. 1334-1345, Aug. 2013.

[5] F. Shao; G. Jiang; W. Lin; M. Yu; Q. Dai, "Joint Bit Allocation and Rate Control for Coding Multi-View Video Plus Depth Based 3D Video," IEEE Transactions on Multimedia, vol. 15, no. 8, pp. 1843-1854, Dec. 2013.

[6] L. Wang; A. Vincent, "Joint rate control for multi-program video coding," International Conference on Consumer Electronics, Digest of Technical Papers, pp. 40-41, June 1996.

[7] L. Boroczky; A. Y. Ngai; E. F. Westermann, "Statistical multiplexing using MPEG-2 video encoders," IBM Journal of Research and Development, vol. 43, no. 4, pp. 511-520, July 1999.

[8] L. Boroczky; A. Y. Ngai; E. F. Westermann, "Joint rate control with look-ahead for multi-program video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, no. 7, pp. 1159-1163, Oct. 2000.

[9] Z. G. Li; C. Zhu; F. Pan; G. Feng; X. Yang; S. Wu; N. Ling, "A novel joint rate control scheme for the coding of multiple real time video programs," International Conference on Distributed Computing Systems Workshops, pp. 241-245, 2002.

[10] J. Yang; X. Fang; H. Xiong, "A joint rate control scheme for H.264 encoding of multiple video sequences," IEEE Transactions on Consumer Electronics, vol. 51, no. 2, pp. 617-623, May 2005.

[11] M. Jacobs; J. Barbarien; S. Tondeur; R. V. Walle; T. Paridaens; P. Schelkens, "Statistical multiplexing using SVC," IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, pp. 1-6, April 2008.

[12] Y. Wang; L. Chau; K. Yap, "Joint Rate Allocation for Multiprogram Video Coding Using FGS," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 6, pp. 829-837, June 2010.

[13] W. Yao; L. Chau; S. Rahardja, "Joint Rate Allocation for Statistical Multiplexing in Video Broadcast Applications," IEEE Transactions on Broadcasting, vol. 58, no. 3, pp. 417-427, Sept. 2012.

[14] W. Yao; L. Chau; S. Rahardja, "Joint Rate Allocation of Stereoscopic 3D Videos in Next-Generation Broadcast Applications," IEEE Transactions on Broadcasting, vol. 59, no. 3, pp. 445-454, Sept. 2013.

[15] Z. He; D. O. Wu, "Linear Rate Control and Optimum Statistical Multiplexing for H.264 Video Broadcast," IEEE Transactions on Multimedia, vol. 10, no. 7, pp. 1237-1249, Nov. 2008.

[16] W. Yao; L. Chau; S. Rahardja, "Joint Rate Allocation For Statistical Multiplexing of SVC," IEEE International Conference on Image Processing, pp. 725-728, Sept. 2012.

[17] H. Fan; L. Ding; X. Xie; H. Jia; W. Gao, "Joint Rate Allocation with Both Look-ahead And Feedback Model For High Efficiency Video Coding," IEEE International Conference on Multimedia and Expo, 2016.

[18] T. Chiang; Y. Zhang, "A new rate control scheme using quadratic rate distortion model," IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no.1, pp. 246-250, Feb. 1997.

[19] G. J. Sullivan; T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Processing Magazine, vol. 15, no. 6, pp. 74-90, Nov. 1998.

[20] S. Mallat; F. Falzon, "Analysis of low bit rate image transform," IEEE Transactions on Signal Processing, vol. 46, no. 4, pp. 1027-1042, April 1998.

[21] M. Dai; D. Loguinov; H. Radha, "Rate-distortion modeling of scalable video coders," International Conference on Image Processing, vol. 2, pp. 1093-1096, Oct. 2004.

[22] M. R. Ardestani; A. A. B. Shirazi; M. R. Hashemi, "Rate-distortion modeling for scalable video coding," IEEE International Conference on Telecommunications, pp. 923-928, April 2010.

[23] B. Li; H. Li; L. Li; J. Zhang, "Lambda Domain Rate Control Algorithm for High Efficiency Video Coding," IEEE Transactions on Image Processing, vol. 23, no. 9, pp. 3841-3854, Sept. 2014.

[24] T. Berger, "Rate Distortion Theory. Englewood Cliffs," NJ: Prentice-Hall, 1971.

[25] T. M. Cover; J. A. Thomas, "Elements of Information Theory", Tsinghua University Press, 1991:155-183.

[26] W. Szepanski, "Δ-entropy and Rate Distortion Bounds for Generalised Gaussian Information Sources and Their Applications to Image Signals", Electronics Letters, vol. 16, no. 3, pp:109-111, 1980

[27] S. Li ; M. Xu ; Z. Wang ; X. Sun, "Optimal Bit Allocation for CTU Level Rate Control in HEVC", IEEE Transactions on Circuits and Systems for Video Technology, vol. PP, no. 99, pp. 1-1, July 2016.

[28] F. Muller, "Distribution shape of two-dimensional DCT coefficients of natural images", Electronics Letters, vol. 29, no. 22, pp.1935-1936, Oct. 1993.

[29] S. Ma; J. Si; S. Wang, "A study on the rate distortion modeling for high efficiency video coding", IEEE International Conference on Image Processing, pp. 181-184, 2012.

[30] Z. Pan; J. Lei, Y. Zhang; X. Sun; S. Kwong, "Fast motion estimation based on content property for low-complexity H.265/HEVC encoder," IEEE Transactions on Broadcasting, vol. 62, no. 3, pp. 675-684, Sept. 2016.

[31] Z. Pan; Y. Zhang; S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," IEEE Transactions on Broadcasting, vol. 61, no. 2, pp. 166-176, 2015.

[32] Z. Pan; P. Jin; J. Lei; Y. Zhang; X. Sun; S. Kwong, "Fast Reference Frame Selection Based on Content Similarity for Low Complexity HEVC Encoder", Journal of Visual Communication and Image Representation, vol. 40, Part B, pp. 516-524, Oct. 2016.

[33] L. Xu; D. Zhao; X. Ji; L. Deng; S. Kwong; W. Gao. "Window level rate control for smooth visual quality and smooth buffer occupancy", IEEE Transactions on Imaging Processing, Vol. 20, no. 3, pp. 723-734, 2011.

[34] Y. Liu; Q. Huang; S. Ma; D. Zhao; W. Gao; S. Ci; H. Tang, "A Novel Rate Control Technique for Multiview Video Plus Depth Based 3D Video Coding", IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 562-571, July 2011.

Hongfei Fan received the B.S. degree in software from Shanghai Jiao Tong University, Shanghai, China, in 2013. From Sep. 2013, he is working toward the Ph.D. degree with the National Engineering Laboratory for Video Technology and the Cooperative Medianet Innovation Center, School of Electronics Engineering and Computer Science, Peking University, Beijing, China. His research interests include video coding and image processing.

Lin Ding received the B.S. degree in computer science from University of Electronic Science and Technology of China, Chengdu, China, in 2013. From Sep. 2013, she is working toward the Ph.D. degree with the National Engineering Laboratory for Video Technology and the Cooperative Medianet Innovation Center, School of Electronics Engineering and Computer Science, Peking University, Beijing, China. Her research interests include video coding and image processing.

Huizhu Jia received his Ph.D. degree in electrical engineering from the Chinese Academy of Sciences, Beijing, China, in 2007. He is currently with the National Engineering Laboratory for Video Technology, Peking University, Beijing. His research interests include: ASIC design, image processing, multimedia data compression and VR.

Xiaodong Xie received his Ph.D. degree in electrical engineering, University of Rochester, USA. He was a Senior Scientist at Eastman Kodak Company, New York, USA, from 1994 to 1997; a Principal Scientist at Broadcom Corporation, California, USA, from 1997 to 2009. He is currently a Professor with the Department of Electrical Engineering, Peking University, Beijing, China. His research interests include multimedia SoC design, embedded system. He holds 24 U.S. Patents.