Available online at www.sciencedirect.com
ScienceDirect
Signal Processing: *Image Communication* 22 (2007) 19–38
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Recent advances in rate control for video coding

Zhenzhong Chen*, King Ngi Ngan

Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Received 8 February 2006; received in revised form 23 August 2006; accepted 2 November 2006

Abstract

In this paper, we review the recent advances in rate control techniques for video coding. The rate control algorithms recommended in the video coding standards are briefly described and analyzed. Recent advances, such as new concepts in rate-distortion modelling and quality constrained control, are presented. With these techniques, the rate control performance can be improved. The paper not only summarizes these recent rate control techniques but also provides explicit directions for future rate control algorithm design.

Keywords: Rate control; Rate distortion; Video coding

1. Introduction

Recent success of multimedia entertainment has generated a lot of research in video coding. To serve the various applications such as video conferencing, Digital versatile disc (DVD), video-on-demand (VOD), digital television broadcasting, Internet video streaming, and digital camcorders, many video coding standards have been developed by the International Telecommunications Union (ITU) and Motion Picture Expert Group (MPEG), e.g. ITU-T H.261, H.263, ISO/IEC MPEG-1 [45], MPEG-2 [46], MPEG-4 [47], and H.264/AVC [54]. These video coding standards employ efficient compression techniques to remove the spatial and temporal redundancy within and between frames. Due to the limited storage size or communication bandwidth, quantization is introduced to reduce the bit rate of the compressed video signal such that the size or bandwidth limitation can be met properly. Rate control mechanism is responsible to adjust the quantization parameters (QPs) to achieve this purpose (Fig. 1).

Generally, the rate control part is an informative part in video coding standards. It leaves the flexibility to designers to develop the suitable scheme for specific applications. Whilst matching the desired bit rate, rate control needs to consider following challenging issues:

(1) *Distortion*: There is an inherent tradeoff between the distortion and the bit rate.
Based on the rate distortion (R-D) theory [3], the distortion D is a decreasing function of the bit-rate R. A decreasing distortion leads to an increasing rate and vice versa. So the fundamental problem in rate control can be

*Corresponding author Tel. +852 26098251; fax: +852 2603 5558
E-mail addresses: zzchen@ieee.org (Z. Chen), knngan@ee.cuhk.edu.hk (K.N. Ngan).

0923-5965/$ - see front matter
doi:10.1016/j.image.2006.11.002

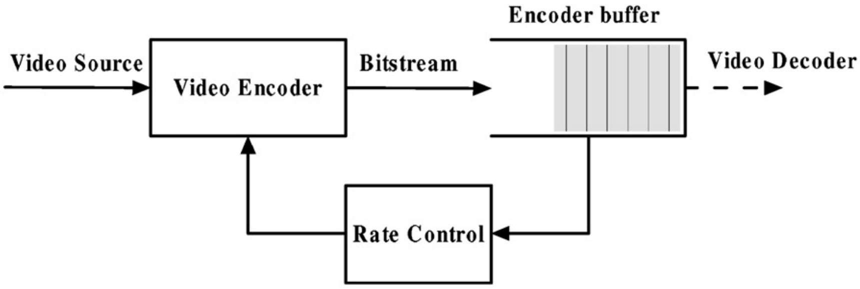

Fig. 1. Rate control in video coding system.

stated as follows:

$$\min \quad D$$
$$\text{s.t.} \quad R \leqslant R_{\max}, \tag{1}$$

where $R_{\max}$ denotes the maximum bit rate. In other words, rate control is to achieve the maximum picture quality (minimum distortion) without exceeding the maximum permitted bit rate where quality is typically represented by peak signal noise ratio (PSNR).

(2) *Complexity*: Different application has different computational complexity requirement. For real-time video coding systems, time consuming rate control algorithm should be avoided. Two-pass rate control algorithm is not applicable for real-time applications while it may be used for storage video. As a dual problem, the rate control algorithm is expected to be simple enough with acceptable performance or is close to the optimal solution with reasonable cost.

(3) *Constraints*: There are various constraints for video compression applications. For end-to-end real-time video communication systems, there is the delay constraint to avoid delay jitter and jerky motion. For constant bit-rate applications, buffer constraint is introduced. The rate control algorithm should guarantee the encoder and decoder buffers do not overflow or underflow. For video storage applications, the budget constraint should be considered since the storage space is fixed.

From the bit-rate point of view, we can classify the video coding applications into two categories: constant bit rate (CBR) and variable bit rate (VBR).

CBR: For CBR applications, rate control designers mainly focus on how to improve the matching accuracy between the target bit rate and actual bit rate and satisfy the low-latency and buffer constraints. In CBR applications, the fluctuation of video quality cannot be avoided due to the varying content in natural scenes. However, by taking the advantage of the encoder buffer, smoothing video quality is possible as the buffer can tolerate limited bit-rate fluctuation whilst the buffer does not either overflow or underflow.

VBR: VBR video coding is employed in some applications in which natural video frames need to be represented with more stable quality. VBR video can also be neatly incorporated in a VBR transmission networking infrastructure [58,75]. For VBR video where delay or rate constraint is not as strict as real-time video coding, VBR rate control is expected to present the video at constant quality for the end users or be adaptive to both the source and channel conditions.

Besides the above issues, different video coding system structure leads to different rate control framework. MPEG-4 object-based video coding system needs joint texture-shape rate control since both texture and shape information can be lossy coded [13]. The rate control algorithm designed for H.264 should consider the flexible block types and the corresponding R-D optimization (RDO) specified in the standard. Recent rate control developments evolve in distortion control. Xie et al. [118] proposed a consistent distortion-based rate control to achieve stable quality in real-time MPEG-4 video coding. Besides, smoothing distortion-based rate control algorithms [39,12] have also been developed to reduce the quality fluctuation between adjacent frames thus making the reconstructed video clips more pleasing.

Generally speaking, the rate control algorithm consists of two main parts: (1) to explore the RD relationship of the video source; (2) to design optimal/suboptimal control techniques to achieve the optimization objects while satisfying the constraints of the applications. Operational R-D (ORD) theory applies to lossy data compression with finite number of possible R-D pairs. These pairs constitute the ORD function [84] and provide sufficient information for optimal rate control or bit allocation. The general objective of optimal rate control is to minimize the distortion under the rate constraint (Eq. (1)). This objective can be further specified as either minimizing average distortion (MINAVE), minimizing maximum distortion (MINMAX) [87] of the video sequence, or minimizing variation (MINVAR) of distortion [66]. Since the generation of the ORD curve by actual R-D pairs has quite high computational complexity and will introduce intolerable delay, the R-D relationships are modelled by the R-D models in many video coding standards or applications. Such a model is expected to be accurate and with limited computational complexity. Many rate control algorithms assume that the source statistics are stationary and belong to some probabilities such as Gaussian [34] or Laplacian [21] and then derive the R-D models based on the R-D theory [21,81]. Meanwhile, with the progress of video coding standardization and explosion of video applications, rate control problem has attracted a lot of attention and been studied extensively. A proper designed rate control approach can efficiently enhance the R-D performance of the video coding system.

In this paper, we give a review of the recent advances in rate control for video coding. The organization of the paper is as follows. We describe the background and preliminaries in next section. A review and discussion of rate control in video coding standards are given in Section 3. In Section 4, the recent advances in rate control algorithm development are then presented. The evolutions in R-D modelling, quality constrained rate control, subjective optimized rate control and other extensions are discussed. We summarize our paper in Section 5.

2. Background and preliminaries

In this section, we review some basic principles related to rate control. We begin with the theoretical foundation of rate control, R-D theory, and the practical concepts, operational and model-based R-D functions. After stating the rate control problems, we will present the commonly used optimization techniques in video coding.

2.1. R-D theory, operational and model-based R-D functions

R-D theory [3] is the theoretical foundation of rate control. It is originated from Shannon's papers [89,90] and forms a central part of information theory [23] and lossy source coding [33,4] which directly relate to lossy image compression [24] and video compression [43]. A lossy source coding scheme, such as video coding, concentrates on the tradeoff between the distortion and bit rate. The basic scenario is a decreasing distortion leads a increasing rate and an increasing distortion is achieved by a decreasing rate. In R-D theory, the R-D function is defined to describe the lower bound for the rate at a given distortion. However, there is no guarantee that the lower bound can be achieved in practical video coding schemes. Comparing with R-D theory, the ORD theory [84] is more applicable for video compression.

ORD theory applies to lossy data compression with finite number of possible R-D pairs. These pairs constitute the ORD function. Detailed definition of these concepts can be found in [84]. A quantizer belongs to the ORD function quantizer set when

$$Q_{\mathrm{ORDF}} = \{q : q \in Q, R(q) \geqslant R(p) \Rightarrow D(q) < D(p),\ \forall p \in Q\}, \quad (2)$$

where $Q = \{q_0, \ldots, q_{M-1}\}$ is the set of all admissible quanitzers, $R(q_i)$ and $D(q_i)$ are the rate and distortion for particular quanitzer, respectively. The ORD function presents the convex curve of the specific compression scheme such that the optimal solution of rate control, i.e., optimal quantizer achieving minimum distortion at given bit rate, can be obtained.

However, the ORD function-based rate control schemes do not work efficiently in many practical video coding applications, especially for the real-time video coding, since the generating of the ORD curve by actual R-D pairs has high computational complexity and introduces intolerable delay. So in many video coding systems, the model-based rate control schemes are adopted. The R-D relationships are modelled by R-D models. Some models are derived based on the statistical properties of video signal and R-D theory [34,21,81], and some models are developed empirically and benefit from various regression techniques [66,26,107,57,9,14]. But these R-D models may suffer from relatively large

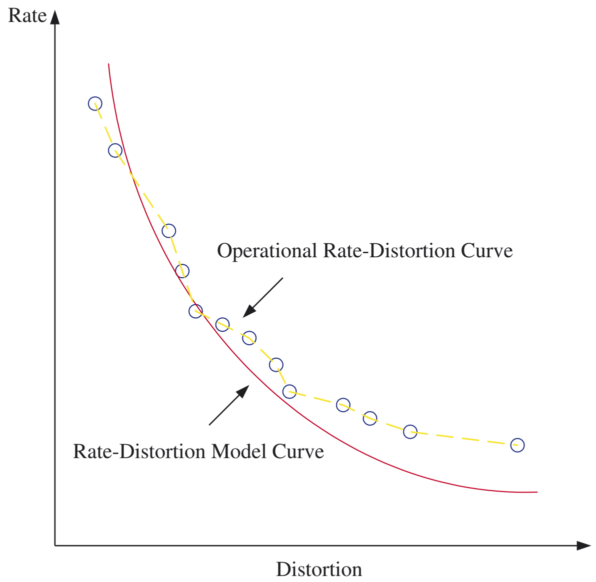

Fig. 2. Operational rate-distortion and rate-distortion model curves.

estimation error as shown in Fig. 2 where the circles are R-D pairs of the admissible quantizers and the solid line shows the approximations by a R-D model. Due to the strict delay or complexity requirements of video coding applications, model-based R-D functions have been widely used in practical video systems. Various expressions of models are applied in video coders and demonstrated great success in applications.

2.2. Optimal rate control problems

The general objective of optimal rate control is to minimize the distortion under the rate constraint (Eq. (1)). This objective can be further specified as either MINAVE or MINMAX [87] of the video sequence. The later one may be extended to lexicographical optimality criterion [40] for near constant distortion. The constant distortion-based rate control problem can be viewed as quality control since it aims to achieve constant distortion and can be further extended to the problem of MINVAR in distortion.

As motion compensated video coding is a dependent process, the bit-rate R_n and the corresponding distortion D_n for particular video frame n depend not only on its QP Q_n, but also on its neighboring frames, i.e., previous a frames and next b frames. The unconstrained MINAVE problem can be formulated as

$$\mathbf{Q}^* = (Q_1^*, \ldots, Q_N^*) = \arg\min_{(Q_1,\ldots,Q_N)} \sum_{n=1}^{N} D_n \times (Q_{n-a}, \ldots, Q_{n+b}), \quad (3)$$

where a and b satisfy the boundary conditions, N presents the number of frames considered in the optimization problem. Accordingly, the MINMAX problem can be formulated as

$$\mathbf{Q}^* = (Q_1^*, \ldots, Q_N^*) = \arg\min_{(Q_1,\ldots,Q_N)} \left(\max_{n\in\{1,\ldots,N\}} \{D_n(Q_{n-a}, \ldots, Q_{n+b})\} \right) \quad (4)$$

and the MINVAR problem can be formulated as

$$\mathbf{Q}^* = (Q_1^*, \ldots, Q_N^*) = \arg\min_{(Q_1,\ldots,Q_N)} \sum_{n=2}^{N} |D_n(Q_{n-a}, \ldots, Q_{n+b}) - D_{n-1}(Q_{n-a-1}, \ldots, Q_{n+b-1})|. \quad (5)$$

It is noted in video coding, the constraints are always applicable. The unconstrained problems (3)–(5) become constrained problems such as delay constrained and buffer constrained. For example, for storage video with a storage space limitation, $R_{\max}$, the MINAVE problem (3) becomes to

$$\mathbf{Q}^* = (Q_1^*, \ldots, Q_N^*) = \arg\min_{(Q_1,\ldots,Q_N)} \sum_{n=1}^{N} D_n \times (Q_{n-a}, \ldots, Q_{n+b}) \quad \text{s.t.} \quad \sum_{n=1}^{N} R_n(Q_{n-a}, \ldots, Q_{n+b}) \leqslant R_{\max}. \quad (6)$$

Similarly, the buffer constrained MINAVE problem can be formulated as

$$\mathbf{Q}^* = (Q_1^*, \ldots, Q_N^*) = \arg\min_{(Q_1,\ldots,Q_N)} \sum_{n=1}^{N} D_n \times (Q_{n-a}, \ldots, Q_{n+b}) \quad \text{s.t.} \quad 0 \leqslant B_n \leqslant B_{\max}, \quad (7)$$

where B_n denotes the buffer level and $B_{\max}$ denotes the maximum allowed buffer level.

It is also noted that for different applications, the constraints are different and multiple constraints may be applicable [74]. How to optimally allocate available bits at frame level or at block level to best represent pictures is the essential problem in rate control. Optimal bit allocation or rate control has been studied extensively in video coding field [76,79,73,74,8,85,86,38,10]. There are two well-known approaches to solve the optimal bit allocation problem, namely Lagrange optimization [27] and dynamic programming (DP) [2]. The optimal bit allocation was first addressed in [42] which used the Lagrange multiplier approach for R-D analysis in transform coding. Further improvements have been reported in [88,92] for source quantization and coding. However, the Lagrange multiplier method may suffer from two main problems, such as having negative bits and real numbers [84]. Further extensions of this method have been reported, i.e., applying the non-negative integer allocation [29] and introducing convex requirement [88]. DP can be employed to achieve the minimum overall distortion through a tree or trellis with known quantizers and their R-D characteristics [76,79]. Particularly, Viterbi algorithm (VA) [28] (or so-called forwarded DP) or shortest path algorithm [22] are widely used. Considering the recursion complexity of DP, some sub-optimal algorithms, such as greedy algorithm [22], are employed in rate control schemes. For the constrained MINAVE and MINVAX problems and solutions, an excellent review can be found in [87].

3. Rate control in video coding standards

Although rate control is not a normative part in any video coding standard, every video coding standard has its own recommendation on rate control as informative part based on the competitions from various related proposals. Considering the balance between the complexity and efficiency, these rate control algorithms do not aim to provide the optimal and complex solutions. The adopted rate control algorithms are competitive in R-D performance with acceptable computational complexities and are flexible in terms of adaptation abilities to various video sources. It is valuable to review the developments of these rate control algorithms in the evolutions of video coding standards since they can reflect the advances achieved in the newly developed rate control techniques. Generally, most rate control algorithms adopted in video coding standards are designed for both CBR and VBR applications. Rate control schemes were developed for the simulation models of the different video coding standards, e.g., reference model (RM) for H.261 [6], adaptive quantization algorithm for MPEG-1 [109], test model (TM) for MPEG-2 [48], test model near-term (TMN) for H.263 [53], verification model (VM) for MPEG-4 [49], and joint model (JM) for H.264 (formerly test model long-term (TML)) [97]. In this section, we review these rate control algorithms recommended in video coding standards.

3.1. H.261

H.261 [52] is a block-based hybrid coder with predictive-DCT transform coding and motion compensation for ISDN video conferencing applications. Since the conference applications are mainly with head-and-shoulder views of people with limited motion and stationary background, the H.261 rate control simply monitors the buffer status to adjust the quantizer step size [6]. The quantizer step size is calculated as a linear function of the buffer level and expressed by

$$q = 2\left\lfloor \frac{B}{200 \times p} \right\rfloor + 2, \tag{8}$$

where p is the multiplier used in specifying the bit rates as in $p \times 64\,\text{kbit/s}$, B is the buffer level, and $\lfloor . \rfloor$ stands for integer division with truncation towards zero. The buffer size B_s is defined as $B_s = p \times 6.4\,\text{kbits}$.

In addition, there are two additional operations in the buffer control to prevent buffer underflow or buffer overflow. If the buffer level reaches the trigger point for overflow, current and subsequent coded data are not sent to allow the buffer to be emptied. On the other hand, bit stuffing is invoked when the buffer is in danger of underflow such that the decoder can maintain synchronization.

3.2. MPEG-1

The MPEG-1 [45] standard is a multimedia standard focused on storage of multimedia content on a standard CD-ROM [35]. The MPEG-1 rate control is to allocate the total number of bits among the various types of pictures (I, P and B-frames). The MPEG video committee suggests that P-pictures be allotted about 2–5 times as many bits as B-pictures, and I-pictures be allotted up to three times as many bits as P-pictures to give good results for typical natural scenes. Under normal circumstances, the encoder monitors the buffer status and adjusts the quantizer step size to avoid both buffer overflow and underflow problems.

In MPEG-1, the adaptive quantization algorithm [109] is recommended. Each MB is classified using class $cl(r,c)$ (coding difficulty), where r, c denote the row and column coordinates of the MB. The quantizer step size for each class $Q(r,c)$ with an overall minimum step size Q_{low} is assigned according to

$$Q(r,c) = Q_{\text{low}} + \Delta Q \times cl(r,c), \tag{9}$$

where ΔQ is typically 1 or 2.

3.3. MPEG-2

MPEG-2 [46] is an extension of MPEG-1 that allows for greater input format flexibility, higher data rates for DVD and high-definition television (HDTV) applications, and better error resilience. The standard was completed in 1994 and is extensively used in the video industry for digital television (DTV) broadcast. The TM5 rate control algorithm [48] was designed for bit-rate control in MPEG-2 video coding standard. It consists of the following steps:

(1) *Target bit allocation*: The target number of bits for the next picture depends on the picture-type and universal weighting factors.
(2) *Rate control*: The reference value of the QP for each macroblock (MB) (Q_j) is set as follows:

$$Q_j = \left(\frac{B_j \times 31}{r}\right), \tag{10}$$

where $r = 2 \times R/f$, R denotes the bit rate, f denotes the frame rate, and B_j is the fullness of the buffer.
(3) *Adaptive quantization*: Finally, the QP for MB j is $mquant_j = Q_j \times N\_act_j$ and is clipped to the range [45,74], where N_act_j is the normalized spatial activity measure for the MB j.

3.4. H.263

H.263 is the first video coding standard for very low-bit-rate video coding. The TMN8 rate control algorithm [53] is designed for H.263 video coding and consists of two levels, frame level and MB level.

(1) *Frame level*: At the frame level, the rate control mechanism determines whether to skip or encode the incoming frame. If the buffer level exceeds a certain threshold, the incoming frame will be skipped. Otherwise, a target number of bits per frame is set based on the fullness of the encoder buffer.
(2) *MB level*: At the MB level, the QP is adapted to achieve the target bit rate. It is based on the following logarithmic R-Q model [81]:

$$R = \begin{cases} \frac{1}{2}\log_2\left(2e^2\frac{\sigma^2}{Q^2}\right) & \text{if } \frac{\sigma^2}{Q^2} > \frac{1}{2e}, \\ \frac{e}{\ln 2}\frac{\sigma^2}{Q^2} & \text{if } \frac{\sigma^2}{Q^2} \leqslant \frac{1}{2e}, \end{cases} \tag{11}$$

where Q is the quantization step size, σ is the standard deviation of the MB. The marcoblock level control is performed for each MB by following steps:

- measure σ,
- calculate the QP Q based on Eq. (11) by given R and σ,
- encode the MB,
- update the model parameters.

3.5. MPEG-4

The intention of the MPEG-4 standard [47] is to support a wider variety of multimedia applications and respective functionalities. Within this context, a scene is viewed as a composition of video objects (VOs) with intrinsic properties such as shape, motion, and texture. Instances of VO in a given time are called video object planes (VOPs). Shape information is binary and is referred to as an alpha plane [5].

MPEG-4 adopted a scalable rate control (SRC) scheme [59] in the VM [49]. This rate control approach consists of following steps:

(1) *Frame level rate control*: The frame rate control in SRC scheme consists of four steps: initialization, calculate target bit rate, compute the QP, updating after encoding. To compute the QP, the quadratic rate-quantizer (R-Q) model is used:

$$R = \frac{X_1 \times S}{Q} + \frac{X_2 \times S}{Q^2}, \tag{12}$$

where X_1 and X_2 denote the model parameters that are updated by linear regression method from previous coded parameters, S is the encoding complexity which is denoted by mean absolute difference (MAD), Q is the QP and R is the target bit rate.
(2) *Multiple VO (MVO) level rate control*: As MPEG-4 supports arbitrarily shaped VO coding, the R-D modelling for MVO is a simple

extension of the above single VO (SVO)[1] R-Q modelling:

$$R_i = \frac{X_{1,i} \times S_i}{Q_i} + \frac{X_{2,i} \times S_i}{Q_i^2}, \tag{13}$$

where i denotes the index of the VO. The target bit rate of the VO depends on its size, motion, and the MAD[2] and the ratio among the vos. The shape rate control is realized through size-conversion. The original binary alpha block (BAB) (16×16 pixel block) can be down-sampled by a conversion ratio (CR) of $\frac{1}{2}$ to 8×8 pixel block or $\frac{1}{4}$ to 4×4 pixel block besides the CR of 0 for full resolution. At the decoder, the data is up-sampled back to the original size. The shape rate control considers the value of maximum tolerable distortion, *AlphaTH*, as the mode decision to decide the suitable conversion ratio.

(3) *MB level rate control*: There are three steps in the flow of MB rate control: target bit-rate estimation, QP calculation, and R-D model update. Essentially, the number of bits R_k^{MB} produced for kth MB is estimated by

$$R_k^{\mathrm{MB}} = \begin{cases} \mathrm{MAD}_k^2 \frac{Y_1}{(Q_k^{\mathrm{MB}})^2} & \textit{if } R_{\mathrm{bpp}}^{\mathrm{VO}} > 0.085, \\ \mathrm{MAD}_k \frac{Y_2}{(Q_k^{\mathrm{MB}})^2} \\ \quad + \mathrm{MAD}_k \frac{Y_3}{(Q_k^{\mathrm{MB}})} & \text{otherwise,} \end{cases} \tag{14}$$

where Y_1, Y_2 and Y_3 are the model parameters, Q_k^{MB} is the MB quantizer, $R_{\mathrm{bpp}}^{\mathrm{VO}}$ denotes the bits per pixels for the texture of the VOP.

3.6. H.264/AVC

Most recently, ITU-T and ISO/IEC developed the newest video coding standard, H.264/advanced video coding (AVC) [116]. It employs some different features such as integer transform, multiple reference frames and various block type-based motion estimation and R-D optimization, and universal variable length coding or context-based adaptive binary arithmetic coding [44]. The rate control algorithm in JM [97] consists of three components:

(1) *GOP level rate control*: GOP level rate control calculates the available bits for the remaining frames in the GOP and initialize the QP of instantaneous decoding refresh (IDR) frame.

(2) *Picture level rate control*: In this level, quantization step Q_{step} is computed by the quadratic model and then used to perform R-D optimization (RDO) [115] for each MB in the frame:

$$R = \frac{a_1 \times \tilde{S}}{Q_{\mathrm{step}}} + \frac{a_2 \times \tilde{S}}{Q_{\mathrm{step}}^2} - h, \tag{15}$$

where h is the number of bits for header and motion vectors, a_1 and a_2 are model parameters estimated by linear regression, and $\tilde{S}$ is the predicted MAD of the current stored picture i which is predicted by a linear model according to the actual MAD of previous stored picture $i-1$:

$$\tilde{S}_i = b_1 \times S_{i-1} + b_2, \tag{16}$$

where b_1 and b_2 are the model parameters updated by linear regression. The corresponding QP Q is then calculated according to Q_{step}.

(3) *Basic unit level rate control*: The basic unit [65] is defined as a picture, slice, MB row or a set of MBs. A linear model is used to predict the MAD of the current basic unit in the frame and a quadratic R-D model is used to calculate the QP which is used for RDO in the basic unit. The basic unit level rate control is to obtain a good tradeoff between the picture quality and bit fluctuation.

3.7. Discussion

From the above descriptions, we can summarize the development of the rate control algorithms in the recent video coding standards as follows:

- R-D models have been introduced since MPEG-4 and H.263. These models are helpful in rate control since they can provide sufficient information for QP decisions. After obtaining the target bit rate for encoding current frame based on the encoder buffer feedback, the quantization can be determined through the R-Q models. The model-based rate control techniques have got great success in video coding systems. In the latest video coding standard, H.264, the model-based rate control is also recommended. But the model is more complicated as the RDO is complementary to rate control. To avoid the expensive iterative solution, the MAD is predicted.

[1] As usual [49], SVO is used to describe the entire frame in MPEG-4. For individual arbitrarily shaped VO, we use individual VO (IVO) in this paper.

Although R-D theory acts as the foundation for R-D modelling, i.e., providing explicit expressions of the R-D functions, the regression analysis and outlier detection techniques provide powerful tools to make the modelling procedure practical.

- More complicated buffer control has evolved to determine the target bit rate of the encoding frame. In the rate control algorithms of earlier video coding standards such as H.261, the QP is updated directly based on the buffer level. In the rate control algorithms of later video coding standards, QPs are not directly dependent on the buffer level since the R-D characteristics of the video signals are available from the R-D models. Several buffer control steps are used to determine the target bit rate. Moreover, in MPEG-4 and H.264, post-encoding process is used to further avoid buffer overflow or underflow as well as to introduce quantizer varying constrains to avoid quality fluctuation.
- Besides H.261, rate control algorithms in video coding standards are adaptive to the video sources by considering their MB-level or frame-level (VOP-level) activities. The characteristics of incoming frames are helpful to build accurate R-D model to provide precise quantizer decision.
- More levels are supported by the rate control algorithms in newer video coding standards, e.g., both MPEG-4 and H.264 support three levels while MPEG-2 and H.263 support two levels. This progress makes the rate control more accurate.

4. Recent advances in rate control

It is noted that the rate control part is informative in video coding standards which means that this part is still open for research to improve after the standard has been finalized. In this section, we review some recent advanced rate control techniques in video coding.

4.1. R-D modelling

In many video coding standards or applications, the R-D relationships are modelled by the R-D models. Such a model is expected to be accurate and with limited complexity. From the classical R-D functions, many mathematical expressions of R-D characteristics have been developed. Many rate control algorithms assume that the source statistics are stationary and belong to some probabilities such as Gaussian [34] or Laplacian [21] and then derive the R-D models based on the R-D theory [110]. Ding et al. [26] proposed a generic R-Q model according to the changes in picture activity and a feedback-based bit allocation. A re-encoding method was developed with the rate-quantization model. In [34], a source model is derived from the classical R-D theory to describe the relationships between the rate, distortion and QPs whilst in [101] an adaptive model-driven bit allocation method based on a parametric R-D model was proposed which incorporated a region classification scheme. Ribas-Corbera et al. [81] presented a logarithmic model for bit rate and distortion and used Lagrange optimization to minimize the distortion subject to the bit-rate constraint. A spline method was reported in [66] and a quadratic R-Q model was proposed in [21], respectively. The quadratic R-Q model was further extended into a scalable rate control algorithm in [59] and optimized for its accuracy in [68]. Cheng et al. [19] studied the linear relationship between the activity measure and bit rate and derived an empirical first-order bits model.

4.1.1. ρ-Domain rate model

Recently, He et al. [36] proposed a linear ρ-domain rate model and the corresponding R-D analysis framework [37]:

$$R(\rho) = \theta(1-\rho), \tag{17}$$

where ρ is measured by the percentage of zeros of the DCT coefficients and θ is a frame constant updated using the linear regression method. The mapping between the QP Q and ρ is calculated based on the distribution of the DCT coefficients. They have shown that in typical transform-based video coding systems, the coding bit rate in the ρ domain can be approximated by a linear function. The model is more accurate than the conventional analytic rate model such that it can achieve target bit rate easier and widely applied as the R-D modelling foundation in some rate control algorithms [39,91].

4.1.2. Separable R-D modelling

In most video coding techniques, the color videos consist of luminance and chrominance components. Fig. 3 shows the luminance (Y) and chrominance (U, V) components of a sample color video, *Mobile*. The Y component provides more texture details while the (U,V) components are more uniform.

Other color video sequences also present similar characteristics. The authors [16] analyzed the R-D model of the color image sequence and investigated the distribution of the integer cosine transform coefficients in H.264 color video coding standards can be approximated by *Laplacian* distribution but the R-D characteristics of the Y and (U,V) components should be modelled separately (Fig. 4). The color video sequence *Mobile* was compressed by H.264 and the R-D curves were plotted in Fig. 4(a). As shown in Fig. 4(a), we can find that the shapes of R-D curves of different components are different due to the different representations of Y and (U,V) textures. Fig. 4(b) shows the bit-rate curves of different components at a fixed QP, $QP = 32$. Although the bit rate of the (U,V) components is lower than that of the Y component, the R-D characteristics of the (U,V) part is different and this difference should not be ignored. This implies that the R-D characteristics of color video signal should be modelled separately. By doing this, the modelling accuracy can be further improved [16]. It should also be noted that we used a 4:2:0 format color video as shown in Fig. 4(b). For other typical color video format such as 4:4:4, the ratio of the (U,V) components in the bit rate would be higher. Although Kamaci et al. [55] proposed to use the Cauchy probability distribution function (PDF)-based R-D modelling, this conclusion still holds. It is obvious that the separable R-D modelling principle is also applicable to other transform-based coding system for color video signal, e.g., DCT-based video coding system such as MPEG1,2,4 and H.26x.

4.1.3. Joint texture shape modelling

One distinct feature of MPEG-4 is that the object-based coding is supported. The scene is viewed as a combination of different VOs. Thus, the coded representation of a VOP includes the representations of its shape, motion and texture. The shape R-D information is very useful to predict the bit rate used for the shape and hence anticipate the amount of rate used for the texture, to assist with buffer stabilization and bit allocation between texture and shape. A distortion prediction is useful to achieve a lower bound on the rate such that the shape information is represented with a given fidelity. A state partition method was proposed in [106] to model the R-D characteristics of binary shapes. In [107], a probabilistic approach was proposed to model the shape based on the statistics of shape information. These two methods were restricted to intra-coded shape only. A linear model for inter-mode shape coding was proposed by the authors [9] which can be extended in rate control for arbitrarily shaped VO coding [11]. In [59,105,82], it was shown how the SVO rate control can be extended to MVO rate control algorithm. The MVO bit allocation problem can be solved by considering the VOs intrinsic properties. Lee et al. [63] considered the rate control issue for different variable frame rates of the VO. Wang et al. [113] applied the ORD-based approach to allocate bits for MVO. Comparing with the traditional frame-based coding, the major difference in the object-based coding is that in addition to coding texture and motion, the shape of the VO must be coded as well. Since both texture and shape can be lossy coded, an object-based rate control scheme is necessary to consider the following two important issues. As the shape information is encoded with the texture information, bit allocation between the texture and shape is the first key problem in MPEG-4 object-based video coding. Another unique issue in MPEG-4 rate control is how to distribute the bit budget for MVOs. A joint

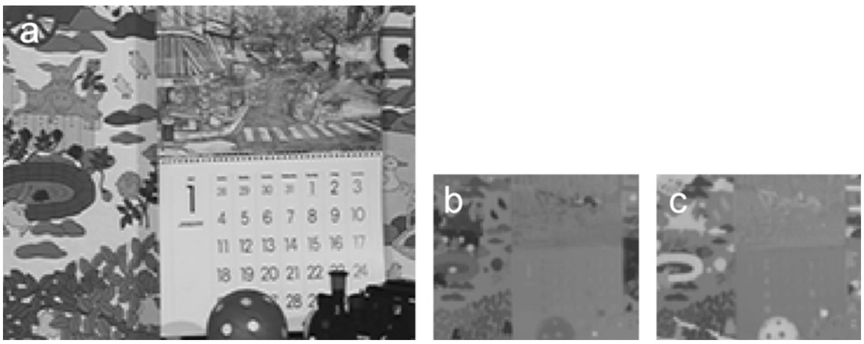

Fig. 3. The luminance and chrominance components of test sequence *Mobile* (QCIF, 4:2:0): (a) Y, (b) U, (c) V.

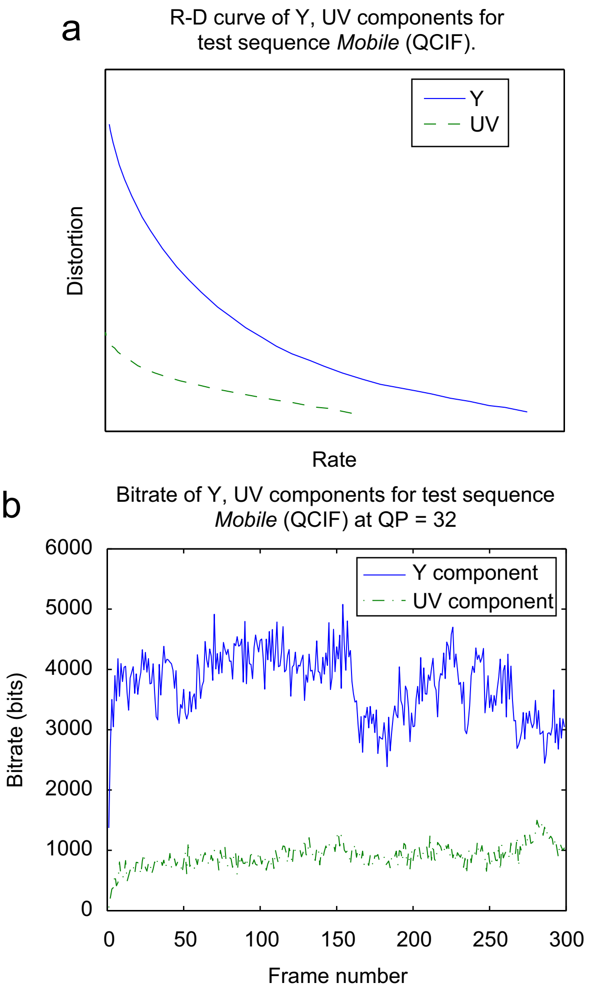

Fig. 4. Test results from test sequence *Mobile* (QCIF, 4:2:0): (a) rate-distortion curves of Y, UV components, (b) bit rates of Y, UV components at $QP = 32$.

texture-shape bit allocation approach was proposed by the authors in [13] to achieve the optimal picture quality. The R-D characteristics of the IVO are obtained via the joint texture-shape R-D model derived from individual texture and shape R-D models. The optimal bit allocation is then performed

$$\min \left(\sum_{i=1}^{m} D_i^{\mathrm{VO}} \right) \quad \text{s.t.} \quad \sum_{i=1}^{m} R_i^{\mathrm{VO}} \leqslant R_{\mathrm{target}}, \tag{18}$$

where D_i^{VO} and R_i^{VO} denote the distortion and bit rate of the ith VO, respectively. m is the number of the VOs.

4.2. Quality constrained rate control

Rate control always evolves into constrained problems in practical applications [77]. Since the amount of information in compressed video sequences is inherently variable, a buffer is placed between the video encoder and the transmission channel to smooth out the rate variation. The buffer also dictates the amount of delay in transmission systems. Larger buffer corresponds to long end-to-end delay. In practical video coding applications, it is essential to consider the rate/buffer/delay constraints. Rate control is then responsible for adjusting the coding parameters to achieve maximum picture quality and ensure the buffer never underflows or overflows.

4.2.1. Constant quality-based rate control

Several rate control techniques have been proposed to achieve constant quality among frames. Overmeire et al. [78] proposed an off-line segment-based rate control approach to maintain constant quality of video sequences. The video is divided into shots based on activity analysis and cut detection. Then the bit allocation is performed in a lexicographic sense among different scenes in a video sequence.

Xie et al. [118] proposed a sequence-based rate control by using the rate-complexity model to achieve consistent quality in real-time MPEG-4 video coding. The sequence-based consistent quality problem was formulated as

$$\min \left(\frac{1}{N} \sum_{i=1}^{N} |D_i(Q_i) - \overline{D(Q)}| \right) \quad \text{s.t.} \quad 0 \leqslant B_i \leqslant B_s, \tag{19}$$

where N is the sequence length, $\overline{D(Q)}$ is the average distortion of all the frames in the sequence, B_s is the encoder buffer size.

4.2.2. Quality smoothing

However, in CBR video coding, the priority is to control the target bit rate without regard to the ensuing visual quality. These strict bit rate control schemes result in large quality fluctuations due to the varying content in natural visual scenes. In contrast to these rate control algorithms, a low-pass filter-based MPEG compression scheme has been proposed by He et al. [39] to smooth the video quality for higher bit-rate applications. They introduced the smoothed rate control framework by using a geometric averaging filter. The low-pass filter was used to smooth the CBR distortion D_{CBR}. In their work, D_{CBR} is estimated by using the linear rate model [36] to solve the constant bit R-Q Q_{CBR} and obtain its corresponding distortion. Then the geometric averaging filter is performed to determine the smoothed distortion D_{S} and its corresponding quantizer Q_{S} for actual coding:

$$D_{\mathrm{S}}(n) = \prod_{i=1}^{N} [D_{\mathrm{CBR}}(n-i)]^{a_i}, \tag{20}$$

where N is the filter length and a_i denotes weighting factors set as $a_i = 1/N$. The quantizer Q_{S} is then decided by the transform coefficients and target distortion D_{S}. This method can smooth out the picture quality fluctuation without sacrificing the overall quality. By combining the buffer control, this low-pass filtering-based approach can achieve smooth quality with limited delay.

As we know, minimizing the averaged distortion may cause the distortion fluctuation problem. In contrast, Lin et al. defined an alternative formation to minimize the distortion variation [66]. The authors [17] proposed a MINVAR of distortion criterion-based approach to reduce the distortion fluctuation between adjacent frames. The distortion variation V is controlled under the bit-rate fluctuation constraint. The distortion variation criterion in the dependent video coding system is defined as [66]

$$\begin{aligned} V_n(Q_{n-a-1}, \ldots, Q_{n+b}) &= |D_n(Q_{n-a}, \ldots, Q_{n+b}) - D_{n-1} \\ &\quad \times (Q_{n-a-1}, \ldots, Q_{n+b-1})|. \end{aligned}$$

So quantizers that satisfy the MINVAR criterion are given by

$$\begin{aligned} \mathbf{Q}^* &= (Q_1^*, \ldots, Q_N^*) \\ &= \underset{(Q_1,\ldots,Q_N)}{\arg\min} \sum_{n=2}^{N} V_n(Q_{n-a-1}, \ldots, Q_{n+b}), \end{aligned} \tag{21}$$

where $n-a-1$ and $n+b$ satisfy the boundary conditions, N presents the number of frames in the sequence. Since in real-time video coding, the coding progress is straightforward such that the dependency of current frame only depended on the coded reference frames. Moreover, the incoming frame can be treated as the non-dependent frame as the previous frames have been coded and will not change their QPs. As encoder buffer allows limited bit-rate fluctuation without buffer overflow or underflow, the MINVAR can be converted to a local optimization problem and can be formulated as a constrained problem from the unconstrained problem:

$$\begin{aligned} Q_n^* &= \underset{Q_n}{\arg\min} \quad V_n(Q_n) \\ \text{s.t.} \quad & R_n^{\mathrm{l}} \leqslant R_n(Q_n) \leqslant R_n^{\mathrm{u}}, \end{aligned} \tag{22}$$

where R_n^{l} and R_n^{u} are the lower and upper bound corresponding to the feedback from the encoder buffer. Fig. 5 shows the comparison of traditional rate control (H.264 JM) and the quality constrained rate control (MINVAR) where the result with smoother PSNR curve, i.e., smoother decoded picture quality, is more pleasing to look at than a fluctuating one. This method can be further extended to the R-D tradeoff framework by combining the buffer variation function [16].

4.3. From objective optimization to subjective optimization

Some recent video coding techniques aim to provide subjectively optimized quality by relying on the characteristics of human visual system [111] or neurobiological properties of visual perception [51].

4.3.1. Foveated video rate control

It is well known that the human visual system (HVS) is highly space-variant and the spatial resolution is highest at the foveation point. So the

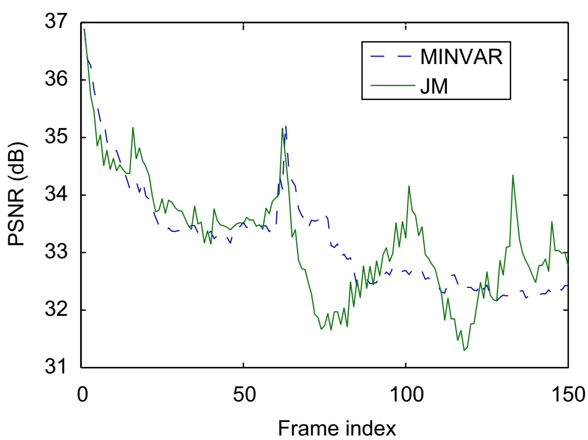

Fig. 5. Comparison of H.264 joint model and MINVAR rate control algorithms for QCIF sequence *Foreman* at bit rate 64 kbps: (a) JM: average PSNR: 33.24 dB, average PSNR variation: 0.27 dB, maximum PSNR variation: 1.24 dB; (b) MINVAR: average PSNR: 33.22 dB, average PSNR variation: 0.11 dB, maximum PSNR variation: 1.15 dB.

object at the foveation point in the scene coded with better quality contributes significantly to the subjective quality. Foveated video compression algorithms [93] have been proposed to deliver high-quality video at reduced bit rates by seeking to match the non-uniform sampling of the human retina. Recently, an optimal rate control approach was developed to maximize a foveal visual quality metric, the foveal signal-to-noise ratio (FSNR), to determine the best compression and rate control parameters for a given target bit rate [62]. The FSNR criterion is defined as an objective quality measurement which matches the non-uniform spatial sensitivity characteristics of human visual system. With this objective criterion, an optimal rate control algorithm employing the Lagrange multiplier method was proposed to maximize the FSNR instead of PSNR which is widely used in traditional coders. Since perceptual quality measurements are complex, an optimal subjective rate control needs to balance the tradeoff between the algorithm complexity and the R-D performance and the perceptual quality can be improved with appropriate quantizer design.

4.3.2. SPEM-based rate control

We have already known that the eye movements have a profound effect on spatio-temporal sensitivity of the human visual system [30], i.e., smooth pursuit eye movements (SPEMs). Nguyen et al. [69] presented an approach based on SPEM, to control the rate of the video data generated by the encoder. The SPEM theory indicates that the eyes always follow the moving object in the visual scene. So Nguyen et al. figured that the VOs with high motion activities should be finely quantized and also found that the high motion part in the video sequence requires more bits for coding. They developed a rate control mechanism to use the SPEM quantizer modifier to improve the quality of moving scenes in the video sequence by using the motion information provided by the motion analysis block together with SPEM component as reference. By adjusting the QP of each frame in the video sequence, the quality of the high motion frames will be improved whilst the quality of those low motion frames will be decreased as a price due to the limited bit budget.

4.3.3. Visual attention-based rate control

With the evolution of computational techniques of visual attention [51], attention properties have been considered in recent video coding applications [50,100,15].

Tang et al. [100] proposed a video bit allocation technique using a visual distortion sensitivity model for H.264 coder. A automatic distortion sensitivity analysis process is employed to analyze both the motion and the texture structures of video frames and evaluate the perceptual distortion sensitivity on a MB basis. The MBs that can tolerate higher perceptual distortion will be allocated fewer bits thus reducing the bit rate of whole frame. Although the objective distortion of the whole frame is higher whilst the bit rate is reduced significantly, no perceptual quality degradation will be perceived by subjective observers.

MPEG-4 treats a scene as a composition of several objects or so-called VOPs that are separately encoded and decoded. Such a flexible video coding framework makes it possible to code different VO with different distortion scale. It is necessary to analyze the priority of the VOs according to its semantic importance, intrinsic properties and psycho-visual characteristics such that the bit budget can be distributed properly to VOs to improve the perceptual quality of the compressed video. Priority-based bit allocation for MVOs in MPEG-4 was proposed in [83]. The VO with higher priority is assumed to be more important and should be encoded with better quality. However, the priority of the VO has to be determined before encoding and cannot be adjusted during encoding. The authors with other colleague [15] introduced an automatic VO priority definition method based on object-level visual attention model and further proposed a subjective optimization framework for VO bit allocation as illustrated in Fig. 6. The human visual system characteristics are incorporated into the video coding optimization process and the priority of the VO can be obtained automatically instead of fixing the weighting factors before encoding or relying on the user interactivity. The authors incorporated the weighting factors of the VOs obtained from the object-level visual attention model with the R-D models to achieve the minimum weighted distortion:

$$\min \quad \left(\sum_{i=1}^{m} W_i D_i^{\mathrm{VO}} \right) \\ \text{s.t.} \quad \sum_{i=1}^{m} R_i^{\mathrm{VO}} \leqslant R_{\mathrm{target}}, \tag{23}$$

where W_i is the weighting factor of the VO i. The total rate for encoding all the VOs, $R(Q_1, \ldots, Q_m)$ is

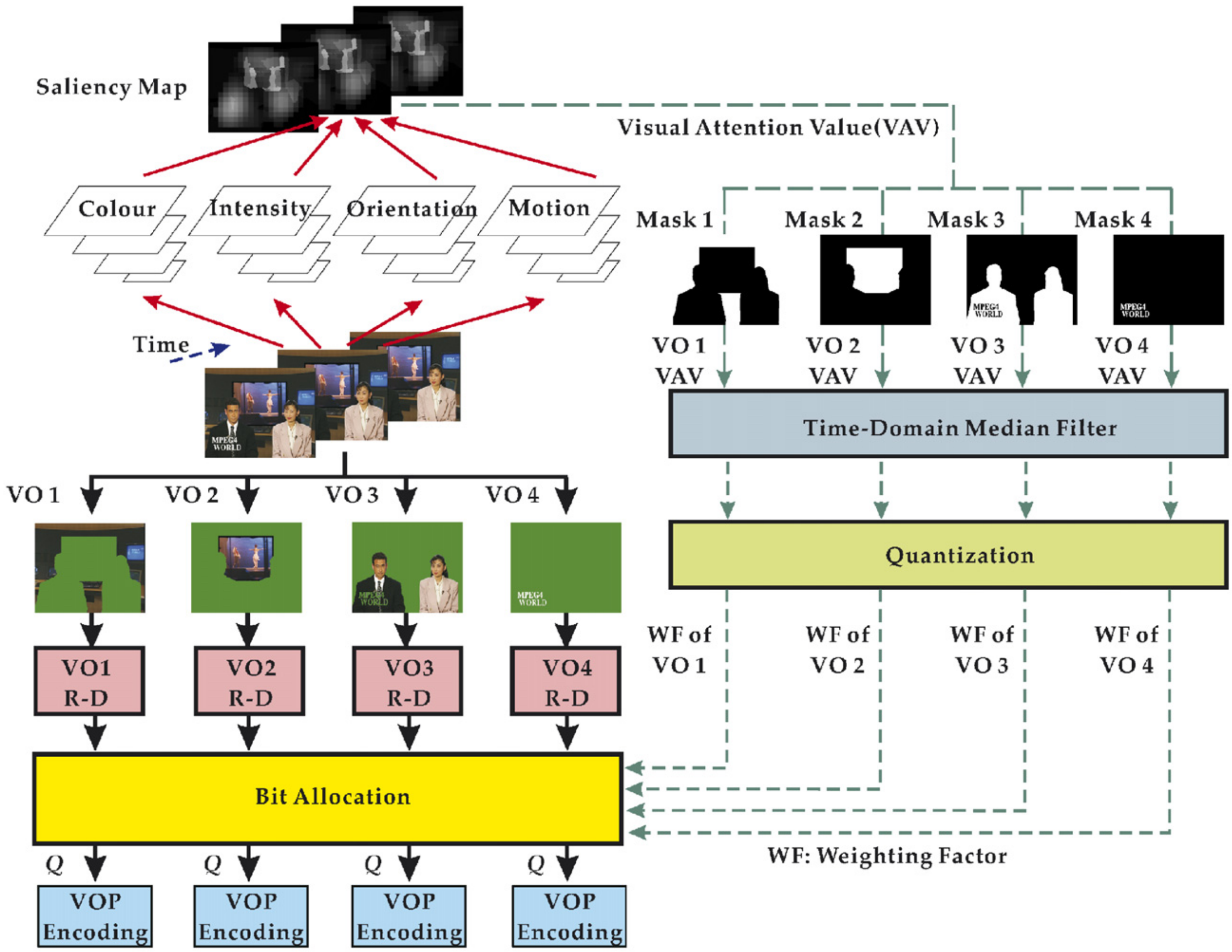

Fig. 6. Framework of unsupervised subjective optimized bit allocation for video coding using visual attention model.

given by the sum of the rates of all the VOs and the distortion for all the VOs, $D(Q_1, \ldots, Q_m)$ is the weighted sum of the distortions of all the VOs. The Lagrange multiplier method can be used to merge the rate and distortion optimizations with a Lagrangian multiplier λ. The constrained problem (23) is converted into the following unconstrained one:

$$\Gamma_\lambda(Q_1, \ldots, Q_m) = D(Q_1, \ldots, Q_m) + \lambda R(Q_1, \ldots, Q_m). \quad (24)$$

It has been shown in [92] that if a λ^* satisfies

$$[Q_1^*, \ldots, Q_m^*] = \arg\min_{Q_1, \ldots, Q_m} \Gamma_{\lambda^*}(Q_1, \ldots, Q_m) \quad (25)$$

and

$$R(Q_1^*, \ldots, Q_m^*) = R_{\text{target}} \quad (26)$$

then $[Q_1^*, \ldots, Q_m^*]$ is the optimal solution of (23). Comparing to the traditional optimal MVO bit allocation algorithm [13], the quality of the object with higher priority (the dancer as shown in Fig. 7) is significantly improved. However, it is noted that the performance relies on the accuracy of the perceptual model.

4.4. Hybrid structure-based optimization

Although various video coding standards have defined common compression architectures of the overall coding systems, the flexibility in the hybrid video coding structure enables the researchers to implement optimization in many bit rate related aspects besides QP such as frame rate, MB mode, and block size.

4.4.1. Mode selection and size selection

In video coding, there are different frame types (I, P, and B frames) and different MB modes (Intra, Inter, skip, etc.). These various types (modes) relate

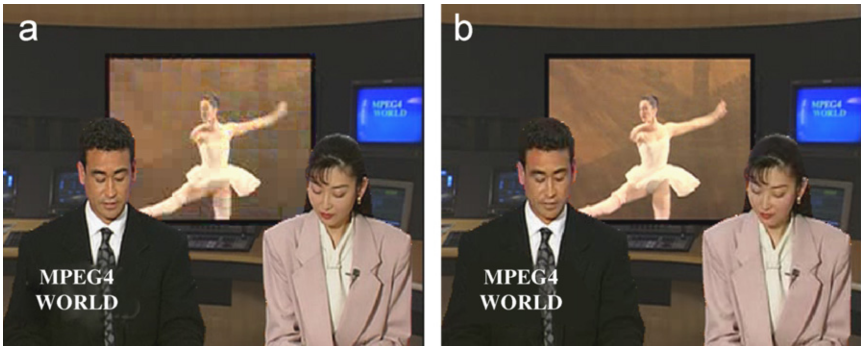

Fig. 7. Reconstructed sample frames of test sequence *News*: (a) optimal bit allocation (average PSNR of VO2: 34.58 dB), (b) subjective optimized bit allocation (average PSNR of VO2: 39.87 dB).

to different bit rates and different residual errors. Frame-type selection can be optimized by incorporating the Lagrange multiplier method to minimize the distortion under the bit budget constraint [60]. The strategy was to determine both the number and the position of reference frames in a group of pictures whilst choosing the optimal quantizer. The mode selection considers the impact of possible modes on R-D behavior in the group of MBs and the optimal strategy is based on the joint selection of MB modes and quantizer scales to achieve the highest picture quality [99]. As H.264 supports flexible block types, e.g., SKIP, INTER16 × 16, INTER16 × 8, INTER8 × 16, INTER8 × 8, INTRA16 × 16, and INTRA4 × 4 for P-slice, the rate control algorithm designed for H.264 consider the block types and the corresponding RDO to determine the coding parameters by using Lagrangian formulation [116]

$$J_M(M|Q, \lambda_M) = D_M(M|Q) + \lambda_M R_M(M|Q), \quad (27)$$

where D_M and R_M are the distortion and bit rate for various modes, respectively. M is the set of modes, Q is the quantizer, and λ_M is the Lagrangian multiplier which has been defined experimentally as

$$\lambda_M = 0.85 \times 2^{(Q-12)/3}. \quad (28)$$

The Lagrange multiplier optimization method provides the optimal mode selection solution. Since there are various mode candidates, the computational complexity of the solution is intensive. To reduce the complexity, some fast mode decision approaches have been proposed by reducing mode candidates [61], joint motion estimation and mode decision [121], or R-D estimation [103].

Besides the flexible mode decision, some rate control algorithms address the problem of flexible decision on block size or frame size. Variable block-size video coding is to balance the tradeoff between the bit rate and distortion by choosing the optimal block size in the frame [95]. Optimal selection of coding parameters in a video coder involves variable-block-size motion estimation and multi-mode residual coding [120]. This size selection is always incorporated in mode selection, especially for H.264.

4.4.2. Frame rate selection

In low bit-rate applications, it is prudent to skip some frames but encode the rest of the frames with better quality. Frame skipping is also widely used to avoid buffer overflow. Song et al. [94] proposed the optimal frame-skipping based on spatial and temporal tradeoff in a R-D sense. Both the frame rate and bit allocation were considered in the rate control framework. They used the mean of the histogram of difference (HOD) image values of current sub-GOP to decide the frame-rate of next sub-GOP. For example, if the mean HOD is larger than a certain threshold, it means that the frame rate could be decreased by one level. In this case, the bits assigned for each coded frame will be higher such that the frame distortion will be smaller. Moreover, a multidimensional rate control algorithm addressed how to jointly control the frame size, frame rate and QPs [80]. In their work, the object function was defined as a weighted distortion metric of coded and skipped frames. The encoding parameters of frame size, frame rate and quantizer were formulated as a buffer constrained bit allocation problem and solved by the DP algorithm.

As MPEG-4 supports MVO coding, different VOs can be encoded at different frame rates. This is essential to reduce the temporal resolution of the stationary VOs such that other objects can gather higher bit rate or higher frame rate to improve the quality. However, the reconstruction mismatches among VOs appear if the boundaries of the VOs vary. Lee et al. [63] proposed a solution by analyzing the shape changes of the VOs and the decoder composition functions such that partial frames of some VOs can be skipped without mismatch impact of the final picture reconstruction while the quality of the overall picture can be improved. The distortions of both coded and skipped VOPs were estimated by R-D models and used for frameskip decision. The composition mismatches were detected and recovered at the decoder for better representation. Sun et al. [98] addressed the problem on how to efficiently code asynchronous objects with different temporal rates so as to achieve a better tradeoff between temporal and spatial resolutions. The VO with higher activity was encoded at higher frame rate and others were encoded at lower frame rate. To deal the composition problem, the method proposed in [63] was adopted.

4.4.3. Multi-level rate control

As we have discussed in Section 3, more levels are supported in rate control for later video coding system. Current rate control works have extended the frame level rate control to MB level. In MPEG-4 VM [49] and H.264/AVC JM [97], MB level or basic unit level rate control strategies have played an important role in the whole rate control system. The multi-level rate control structure makes the rate control more accurate and robust. It is obvious that the MINAVE, MINMAX or MINVAR problems can be formulated in MB levels and the optimal bit allocation among MBs for certain objective can be achieved through the optimization techniques as we have discussed in Section 2. However, we are also aware of the high computational complexity of the optimization problem and notice that some practical solutions have been provided to achieve good R-D performance.

He et al. proposed a solution for MB level QP adjustment besides the GOP and frame level rate control [36]. They noticed that the frame level QP solved by the target bit rate and the R-Q model is a real number but the coder only support integer quantizer parameters. If we simply round the value, the accuracy of the rate control will be decreased. They proposed two solutions. The first one is to arbitrarily let some of MB quantizers be the largest integer smaller than the real number and others be the smallest integer larger than the real number whilst the mean quantization value equals to the real number. More reasonable, they suggested another MB feedback solution. They considered the number of percentage of zeros in the MB and define threshold to determine the quantizer value. This approach is also simple but performs well.

Chai et al. [7] proposed a foreground/background video coding scheme by segmenting the frame as foreground and background regions and then encoding different regions using the same coder but with different quantization step sizes to improve the subjective quality of the decoded video sequences. In this scheme, rate control not only works for frame level but also at sub-region level. The idea behind this method is that people are more interested in the foreground area and improving the objective quality of foreground object will improve the subjective quality of the visual scene. Recently, with the progress of visual attention research, we know that different MB has different visual property. Tang et al. [100] considered the visual distortion sensitivity of MB and allocated fewer bits (larger quantizer) to MBs that can tolerate higher perceptual distortion thus reducing the bit rate of the whole frame. This algorithm can reduce the bit rate without causing perceptual quality degradation.

Besides the above aspects, we are aware of the variety of rate control algorithms in the literature. Most rate control algorithms adjust the QPs based on the buffer feedback and the R-D model. A post-quantizer control technique was proposed by the authors in [68] which jointly considered the variations of the buffer level and the texture complexity measurement to avoid buffer overflow and underflow as well as optimize the picture quality. The function of rate control algorithm in VBR video coding aims to determine the QP to maximize consistent video quality. Besides one-pass rate control algorithms [67] which are suitable for live VBR video encoding, two-pass VBR rate control algorithms can be developed for storage applications [114] which are strictly under budget constraint instead of delay constraint. Using the two-pass approach, the encoder is able to gather the R-D information of the entire video sequence such that the encoder can control quality smoothly and the bit budget precisely during the second-pass coding [122].

5. Summary

Most video coding standards do not specify the rate control mechanism of the encoder as a normative part. The implementation of rate control algorithm is left open to the designers to develop based on the application requirements. As a key component in video coding, rate control has been studied extensively as reviewed in this paper. As the ORD theory [84] setups the framework for optimal solution of rate control, i.e., optimal quantizer achieving minimum distortion at given bit rate, model-based rate control schemes [13,66,21,26, 105,123] approximate the optimal solution and reduce computational complexities for practical applications [43]. Based on the R-D characteristics, many techniques have been developed to achieve balance among efficiency and complexity, e.g., mode selection, and post-quantizer control. We realize the challenges and contributions of the rate control on how to improve the subjective quality of the video by improving the objective quality of human interested object (region) [15]. Achieving smooth video quality in real-time video coding is also an interesting topic in the recent rate control algorithm design [39,12]. Whilst new video coding standards or techniques, such as H.264 [96], wavelet video coding [72], and distributed video coding [32], provide new modes and functions, it is challenging to develop new adaptive rate control algorithms.

In this paper, we summarized recent advances in rate control developments. These new techniques can be carefully selected and applied in different applications. Since different video coding systems have different characteristics, how to define a proper rate control problem is crucial for rate control design. As we have mentioned, the video coding system have determined how many levels the rate control can support. For example, in H.264, the rate control could be effective in GOP, frame and basic unit levels. Besides, in the multi-view video [71] applications, we not only need to solve the rate control problem in a single-view coder, but also need to consider this problem for all views simultaneously. Generally speaking, the video coding system structure and application background play an very important role in rate control algorithm design. It is also noted that rate control also plays an important role in video transmissions over wired [108] wireless networks [18]. Although this rate control is a transport approach instead of video source oriented, we are aware that the proper rate control can significantly improve the system performance by reducing time-out effects, packet loss thus enhance the video quality and guarantee quality of service (QoS) [117,70].

In this paper, we mainly review the rate control techniques for video source coding. Rate control is still a challenging problem in many aspects. We highlight some future research directions:

(1) *Perceptually friendly video coding*: Some video coding researchers aim to develop perceptually friendly coding schemes, e.g. foveated video compression [93], to deliver high-quality video at reduced bit rates by seeking to match the non-uniform sampling of the human retina. An optimal rate control approach is then to maximize a foveal visual quality metric by determine the best rate control parameters for a given target bit rate [62]. Since perceptual quality measurement is very complex, optimal subjective rate control needs to balance the tradeoff between the algorithm complexity and the R-D performance.

(2) *Scalable video coding*: Scalable video coding is crucial in the heterogeneous environment due to its adaptivity to varying network bandwidth and conditions [31] and simultaneously support for both constant bit rate and constant distortion service [102,64]. The R-D characteristics can be extracted from the layered signals and be used to determine the rate-allocation scheme to adaptively achieve minimum quality variation [124,112].

(3) *Joint source-channel coding*: Subject to the constraints in communication systems, rate control in joint source-channel coding is to find the optimal distribution of bits between the source coding and channel coding to achieve the minimum end-to-end distortion. A joint encoder and channel rate control for variable bit-rate (VBR) video over packet-switched networks should be able to balance both the consistent video quality on the encoder side and bitstream smoothness on the networks side [25]. For constant bit rate (CBR) environment, rate control optimizes the quality of the transmitted video by jointly selecting the source rate and the channel rate subject to two sets of constraints, the end-to-end delay and the consistent transmission rate [41,20].

(4) *Transcoding*: Digital video applications benefit from the standardization of the bit stream format

and coding methods as in the video coding standards. Video transcoding [56,104,119] is a conversion process from a compressed video with a certain data rate to one with a different rate or between different video standards, e.g., from MPEG-2 to MPEG-4. Rate control is one of the key functions in a transcoder for rate-adaptation and bitstream concatenation and directly influences the video quality degradation introduced by the transcoding process [1].

We focus on, but not limited to, the above areas for further study on rate control extensions and are aware of the importance of rate control in video coding applications. In this paper, we investigated the contributions of R-D theory in video coding. We demonstrated with different video coding applications that appropriate rate-control techniques lead to successful implementations with superior performances.

Acknowledgment

This work is affiliated with the Microsoft-CUHK Joint Laboratory for Human-centric Computing and Interface Technologies.

References

[1] P.A.A. Assunção, M. Ghanbari, A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams, IEEE Trans. Circuits Systems Video Technol. 8 (1998) 953–967.
[2] R.E. Bellman, Dynamic Programming, Princeton University Press, Princeton, NJ, 1957.
[3] T. Berger, Rate Distortion Theory, Prentice-Hall, Englewood Cliffs, NJ, 1971.
[4] T. Berger, J.D. Gibson, Lossy source coding, IEEE Trans. on Inform. Theory 44 (1998) 2693–2723.
[5] N. Brady, MPEG-4 standardized methods for the compression of arbitrarily shaped video objects, IEEE Trans. Circuits Systems Video Technol. 9 (1999) 1170–1189.
[6] CCITT SG XV WP/1/Q4, Description of reference model 8 (RM8), June 1989.
[7] D. Chai, K.N. Ngan, Foreground/background video coding scheme, in: Proceedings of IEEE International Conference on Circuits and Systems, ISCAS '97, Hong Kong, 1997.
[8] J.-J. Chen, D.W. Lin, Optimal bit allocation for video coding under multiple constraints, in: Proceedings of the IEEE International Conference Image Processing, Lausanne, Switzerland, 1996.
[9] Z. Chen, K.N. Ngan, Linear rate-distortion models for MPEG-4 shape coding, IEEE Trans. Circuits Systems Video Technol. 14 (2004) 869–873.
[10] Z. Chen, K.N. Ngan, Optimal bit allocation for MPEG-4 multiple video objects, in: Proceedings of the IEEE International Conference Image Processing, Singapore, 2004.
[11] Z. Chen, K.N. Ngan, Rate-constrained arbitrarily shaped video object coding with object-based rate control, IEE-Vision Image Signal Process. 151 (2004) 250–256.
[12] Z. Chen, K.N. Ngan, Minvar: a local optimization criterion for rate-distortion tradeoff in real time video coding, in: Proc. SPIE on Multimedia Systems and Applications VIII, Boston, MA, 2005.
[13] Z. Chen, K.N. Ngan, Joint texture-shape optimization for MPEG-4 multiple video objects, IEEE Trans. Circuits Systems Video Technol. 15 (2005) 1170–1174.
[14] Z. Chen, K.N. Ngan, Rate-distortion analysis for MPEG-4 binary shape coding, in: IEEE International Symposium on Intelligent Signal Processing and Communications Systems, Hong Kong, 2005.
[15] Z. Chen, K.N. Ngan, A unified framework of unsupervised subjective optimized bit allocation for video coding using visual attention model, in: Proceedings of the SPIE on Multimedia Systems and Applications VIII, Boston, MA, 2005.
[16] Z. Chen, K.N. Ngan, Towards R-D tradeoff in real-time color video coding, in: IEEE International Symposium on Circuit and System, Koc, Greece, 2006.
[17] Z. Chen, K.N. Ngan, Distortion variation minimization in real-time video coding, Signal Processing: Image Communications 21 (2006) 273–279.
[18] M. Chen, A. Zakhor, Rate control for streaming video over wireless, IEEE Wireless Comm. 12 (2005) 32–41.
[19] J.-B. Cheng, H.-M. Hang, Adaptive piecewise linear bits estimation model for MPEG based video coding, J. Visual Comm. Image Representation 8 (1997) 51–67.
[20] G. Cheung, A. Zakhor, Bit allocation for joint source/channel coding of scalable video, IEEE Trans. Image Process. 9 (2000) 340–356.
[21] T. Chiang, Y.-Q. Zhang, A new rate control scheme using quadratic rate-distortion modeling, IEEE Trans. Circuits Systems Video Technol. 7 (1997) 246–250.
[22] T. Cormen, C. Leiserson, R. Rivest, Introduction to Algorithms, MIT Press, Cambridge, MA, 1990.
[23] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[24] L.D. Davisson, Rate-distortion theory and applications, Proc. IEEE 60 (1972) 800–808.
[25] W. Ding, Joint encoder and channel rate control of VBR video over ATM networks, IEEE Trans. Circuits Systems Video Technol. 7 (1997) 266–2278.
[26] W. Ding, B. Liu, Rate control of MPEG video coding and recording by rate-quantization modeling, IEEE Trans. Circuits Systems Video Technol. 6 (1996) 12–20.
[27] H. Everett, Generlized Lagrange multiplier method for solving problems of optimum allocation of resource, Oper. Res. 11 (1963) 399–417.
[28] G.D. Forney, The Viterbi algorithm, Proc. IEEE 61 (1973) 268–278.
[29] B. Fox, Discrete optimization via marginal analysis, Management Sci. 13 (1966) 210–216.
[30] B. Girod, Eye movements and coding of video sequences, in: SPIE Visual Communications and Image Processing, 1988, pp. 398–405.

[31] M. Ghanbari, Standard Codecs: Image Compression to Advanced Video Coding, IEE Press, New York, 2003.

[32] B. Girod, A.M. Aaron, S.R.D. Rebollo-Monedero, Distributed video coding, Proc. IEEE 93 (2005) 71–83.

[33] R.M. Gray, Source Coding Theory, Kluwer, Norwell, MA, 1990.

[34] H.-M. Hang, J.J. Chen, Source model for transform video coder and its application-Part I: fundamental theory, IEEE Trans. Circuits Systems Video Technol. 7 (1997) 287–298.

[35] B.G. Haskell, P.G. Howard, Y.A. LeCun, A. Puri, J. Ostermann, M.R. Civanlar, L. Rabiner, L. Bottou, P. Haffner, Image and video coding—emerging standards and beyond, IEEE Trans. Circuits Systems Video Technol. 8 (1998) 814–837.

[36] Z. He, Y.K. Kim, S.K. Mitra, Low-delay rate control for DCT video coding via ρ domain source modeling, IEEE Trans. Circuits Systems Video Technol. 11 (2001) 928–940.

[37] Z. He, S.K. Mitra, A unified rate-distortion analysis framework for transform coding, IEEE Trans. Circuits Systems Video Technol. 11 (2001) 1221–1236.

[38] Z. He, S.K. Mitra, Optimum bit allocation and accurate rate control for video coding via-domain source modeling, IEEE Trans. Circuits Systems Video Technol. 12 (2002) 840–849.

[39] Z. He, W. Zeng, C.W. Chen, Lowpass filtering of rate-distortion functions for quality smoothing in real-time video communication, IEEE Trans. Circuits Systems Video Technol. (2005) 973–981.

[40] D.T.Y. Hoang, E.L. Linzer, J.S. Vitter, Lexicographic bit allocation for MPEG video, J. Visual Comm. Image Representation 8 (1997) 384–404.

[41] C.-Y. Hsu, A. Ortega, A.R. Reibman, Joint selection of source and channel rate for VBR vieo transmission under ATM policing constraints, IEEE J. Select. Areas Comm. 15 (1997) 1016–1028.

[42] J.J.Y. Huang, P.M. Schultheiss, Block quantization of correlated Gaussian random variables, IEEE Trans. Comm. System 11 (1963) 289–296.

[43] IEEE Signal Processing Magazine, 1998.

[44] IEEE Trans. Circuits Syst. Video Technol., special issue on the H.264/AVC video coding standard, vol. 13, 2003.

[45] ISO/IEC 11172-2:1991, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbps, Part 2: Visual, 1991.

[46] ISO/IEC 13818-2:1994, Information technology—generic coding of moving pictures and associated audio, Part 2: Visual, 1994.

[47] ISO/IEC 14496-2:1999, Information technology—coding of audio/visual objects, Part 2: Visual, 1999.

[48] ISO/IEC JTC/SC29/WG11, MPEG test model 5, April 1993.

[49] ISO/IEC JTC1/SC29/WG11, MPEG-4 video verification model v18.0, January 2001.

[50] L. Itti, Automatic foveation for video compression using a neurobiological model of visual attention, IEEE Trans. Image Process. 13 (2004) 1304–1318.

[51] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Machine Intell. 20 (1998) 1254–1259.

[52] ITU-T Recommendation H.261 version 2, Video codec for audiovisual services at $p \times 64\,\mathrm{kbit/s}$, 1990.

[53] ITU-T/SG15, Video codec test model, TMN8, June 1997.

[54] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264|ISO/IEC 14496-10 AVC), JVT-G050, 2003.

[55] N. Kamaci, Y. Altunbasak, R.M. Mersereau, Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models, IEEE Trans. Circuits Systems Video Technol. (2005) 994–1006.

[56] G. Keesman, R. Hellinghuizen, F. Hoeksema, G. Heideman, Transcoding MPEG bitsteams, Signal Processing: Image Communications 8 (1996) 481–500.

[57] H.M. Kim, Adaptive rate control using nonlinear regression, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 432–439.

[58] T.V. Lakshman, A. Ortega, A.R. Reibman, VBR video: tradeoffs and potentials, Proc. IEEE 86 (1998) 952–973.

[59] H.-J. Lee, T. Chiang, Y.-Q. Zhang, Scalable rate control for MPEG-4 video, IEEE Trans. Circuits Systems Video Technol. 10 (2000) 878–894.

[60] J. Lee, B.W. Dickinson, Rate-distortion optimized frame type selection for MPEG encoding, IEEE Trans. Circuits Systems Video Technol. 7 (1997) 501–510.

[61] J. Lee, B. Jeon, Fast mode decision for H.264, in: Proceedings of IEEE International Conference on Multimedia and Expo, Taipei, Taiwan, 2004.

[62] S. Lee, M.S. Pattichis, A.C. Bovik, Foveated video compression with optimal rate control, IEEE Trans. Circuits Systems Video Technol. 10 (2001) 977–992.

[63] J.-W. Lee, A. Vetro, Y. Wang, Y.-S. Ho, Bit allocation for MPEG-4 video coding with spatio-temporal tradeoffs, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 488–502.

[64] W. Li, Overview of fine granularity scalability in MPEG-4 video standard, IEEE Trans. Circuits Systems Video Technol. 11 (2001) 301–317.

[65] Z. Li, F. Pan, K.-P. Lim, G. Feng, X. Lin, S. Rahardja, Adaptive basic unit layer rate control for JVT, JVT-G012, March 2003.

[66] L.-J. Lin, A. Ortega, Bit-rate control using piecewise approximated rate-distortion characteristics, IEEE Trans. Circuits Systems Video Technol. 8 (1998) 446–459.

[67] N. Mohsenian, R. Rajagopalan, C.A. Gonzales, Single-pass constant- and variable-bit-rate MPEG-2 video compression, IBM J. Res. Develop. 43 (1999) 489–509.

[68] K.N. Ngan, T. Meier, Z. Chen, Improved single video object rate control for MPEG-4, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 385–393.

[69] A.G. Nguyen, J.-N. Hwang, A novel hybrid HVPC/mathematical model rate control for low bit-rate streaming video, Signal Processing: Image Communications 17 (2002) 423–440.

[70] Z. Ni, Z. Chen, K.N. Ngan, A real-time video transport system for the best-effort Internet, Signal Processing: Image Communications 20 (2005) 277–293.

[71] J.-R. Ohm, Stereo/multiview encoding using the MPEG family of standards, in: Proceedings of SPIE Conference on Electronic Imaging, EI'97, San Diego, USA, 1999.

[72] J.-R. Ohm, Advances in scalable video coding, Proc. IEEE 93 (2005) 42–56.

[73] A. Ortega, Optimization techniques for adaptive quantization of image and video under delay constraints, Ph.D. Thesis, Columbia University, New York, 1994.

[74] A. Ortega, Optimal bit allocation under multiple rate constraints, in: Proceedings of the Data Compression Conference, Snowbird, UT, 1996, pp. 349–358.
[75] A. Ortega, Variable bit-rate video coding, in: M.-T. Sun, A.R. Reibman (Eds.), Compressed Video over Networks, Marcel Dekker, New York, NY, 2000, pp. 343–382.
[76] A. Ortega, K. Ramchandran, M. Vetterli, Optimal trellis-based buffered compression and fast approximations, IEEE Trans. Image Process. 3 (1994) 26–40.
[77] A. Ortega, K. Ramchandran, Rate-distrotion methods for image and video compression, IEEE Signal Process. Magazine (1998) 23–50.
[78] L. Overmeire, L. Nachtergaele, F. Verdicchio, J. Barbarien, P. Schelkens, Constant quality video coding using video content analysis, Signal Processing: Image Communications 20 (2005) 343–369.
[79] K. Ramchandran, A. Ortega, M. Vetterli, Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders, IEEE Trans. Image Process. 3 (1994) 533–545.
[80] E.C. Reed, J.S. Lim, Optimal multidimensional bit-rate control for video communications, IEEE Trans. Image Process. 11 (2002) 873–885.
[81] J. Ribas-Corbera, S. Lei, Rate control in DCT video coding for low-delay communications, IEEE Trans. Circuits Systems Video Technol. 9 (1999) 172–185.
[82] J. Ronda, M. Eckert, F. Jaureguizar, N. Garcia, Rate control and bit allocation for MPEG-4, IEEE Trans. Circuits Systems Video Technol. 9 (1999) 1243–1258.
[83] J.I. Ronda, M. Eckert, S. Rienke, F. Jaureguizar, A. Pacheco, Advanced rate control for MPEG-4 coders, in: Proceedings of the Visual Communications and Image Processing, San Jose, CA, 1998, pp. 383–394.
[84] G.M. Schuster, A.K. Katsaggelos, Rate Distortion based Video Compression, Kluwer Academic Publishers, Norwell, MA, 1997.
[85] G.M. Schuster, A.K. Katsaggelos, A video compression scheme with optimal bit allocation among segmentation motion and residual error, IEEE Trans. Image Process. 6 (1997) 1487–1502.
[86] G.M. Schuster, A.K. Katsaggelos, A theory for the optimal bit allocation between displacement vector field and displaced frame difference, IEEE J. Select. Areas Comm. 15 (1997) 1739–1751.
[87] G.M. Schuster, G. Melnikov, A.K. Katsaggelos, A review of the minimum maximum criterion for optimal bit allocation among dependent quantizers, IEEE Trans. Multimedia 1 (1999) 3–17.
[88] A. Segall, Bit allocation and encoding for vector sources, IEEE Trans. Inform. Theory 22 (1976) 162–169.
[89] C.E. Shannon, A mathematical theory of communications, Bell System Technol. J. 27 (1948) 397–423 623–656.
[90] C.E. Shannon, Coding theorems for a discrete source with a fidelity criterion, in: IRE National Convention Record, vol. 4, New York, NY, 1959, pp. 142–163.
[91] I.-H. Shin, Y.-L. Lee, H. Park, Rate control using linear rate-ρ model for H.264, Signal Processing: Image Communications 19 (2004) 441–452.
[92] Y. Shoham, A. Gersho, Efficient bit allocation for an arbitrary set of quantizers, IEEE Trans. Acoust. Speech Signal Process. 36 (1988) 1445–1453.
[93] P.L. Silsbee, A.C. Bovik, D. Chen, Visual pattern image sequence coding, IEEE Trans. Circuits Systems Video Technol. 3 (1993) 291–301.
[94] H. Song, C.-C.J. Kuo, Rate control for low-bit-rate video via variable-encoding frame rates, IEEE Trans. Circuits Systems Video Technol. 11 (2001) 512–521.
[95] G.J. Sullivan, R.L. Baker, Efficient quadtree coding of images and video, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Singal Processing, Toronto, Canada, 1991.
[96] G.J. Sullivan, T. Wiegand, Video compression—from concepts to the H.264/AVC standard, Proc. IEEE 93 (2005) 18–31.
[97] G. Sullivan, T. Wiegand, K.-P. Lim, Joint model reference encoding methods and decoding concealment methods, JVT-I049, September 2003.
[98] Y. Sun, I. Ahmad, Asynchronous rate control for multi-object videos, IEEE Trans. Circuits Systems Video Technol. 15 (2005) 1007–1018.
[99] H. Sun, W. Kwok, M. Chien, C.H.J. Ju, MPEG coding performance improvement by jointly optimizing coding mode decisions and rate control, IEEE Trans. Circuits Systems Video Technol. 7 (1997) 449–458.
[100] C.-W. Tang, C.-H. Chen, Y.-H. Yu, C.-J. Tsai, Visual sensitivity guided bit allocation for video coding, IEEE Trans. Multimedia 8 (2006) 11–18.
[101] B. Tao, B.W. Dickinson, H.A. Peterson, Adaptive model-driven bit allocation for MPEG video coding, IEEE Trans. Circuits Systems Video Technol. 10 (2000) 147–157.
[102] D. Taubman, A. Zakhor, A common framework for rate and distortion based scaling of highly scalable compressed video, IEEE Trans. Circuits Systems Video Technol. 6 (1996) 329–354.
[103] Y.-K. Tu, J.-F. Yang, M.-T. Sun, Rate-distortion estimation for H.264/AVC coders, in: Proceedings of IEEE International Conference on Multimedia and Expo, Amsterdam, Netherlands, 2005.
[104] A. Vetro, C. Christopoulos, H. Sun, An overview of video transcoding architectures and techniques, IEEE Signal Process. Magazine (2003) 18–29.
[105] A. Vetro, H. Sun, Y. Wang, MPEG-4 rate control for multiple video objects, IEEE Trans. Circuits Systems Video Technol. 9 (1999) 186–199.
[106] A. Vetro, H. Sun, Y. Wang, O. Guleryuz, Rate-distortion modeling of binary shape using state partitioning, in: Proceedings of the IEEE International Conference on Image Processing, Kobe, Japan, 1999.
[107] A. Vetro, Y. Wang, H. Sun, A probabilistic approach for rate-distortion modeling of multiscale binary shape, IEEE Trans. Image Process. 12 (2003) 356–364.
[108] J. Viéron, C. Guillemot, Real-time constrained TCP-compatible rate control for video over the Internet, IEEE Trans. Multimedia 6 (2004) 634–646.
[109] E. Viscito, C. Gonzales, A video compression algorithm with adaptive bit allocation and quantization, in: Proceedings of the Visual Communications and Image Processing (VCIP'91), vol. 1605, SPIE, Boston, MA, 1991, pp. 58–72.
[110] A.J. Viterbi, J.K. Omura, Principle of Digital Communication and Coding, McGraw-Hill, New York, 1979.
[111] Z. Wang, L. Lu, A.C. Bovik, Foveation scalable video coding with automatic fixation selection, IEEE Trans. Image Process. 12 (2003) 243–254.

[112] Y. Wang, J. Ostermann,1 Y.-Q. Zhang, Video Processing and Communications, Prentice-Hall, Englewood Cliffs, NJ, 2001.

[113] H. Wang, G.M. Schuster, A.K. Katsaggelos, Rate-distortion optimal bit allocation scheme for object-based video coding, IEEE Trans. Circuits Systems Video Technol.

[114] P.H. Westerink, R. Rajagopalan, C.A. Gonzales, Two-pass MPEG-2 variable-bit-rate encoding, IBM J. Res. Develop. 43 (1999) 471–488.

[115] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, G.J. Sullivan, Rate-constrained coder control and comparison of video coding standards, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 688–703.

[116] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, A. Luthra, Overview of the H.264/AVC video coding standard, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 560–576.

[117] D. Wu, T. Hou, Y.-Q. Zhang, Transporting real-time video over the Internet: challenges and approaches, Proc. IEEE 88 (2000) 1855–1875.

[118] B. Xie, W. Zeng, A sequence-based rate control framework for consistent quality real-time video, IEEE Trans. Circuits Systems Video Technol. 16 (2006) 56–71.

[119] J. Xin, C.-W. Lin, M.-T. Sun, Digital video transcoding, Proc. IEEE 93 (2005) 84–97.

[120] Y. Yang, S.S. Hemami, Generalized rate-distortion optimization for motion-compensated video coders, IEEE Trans. Circuits Systems Video Technol. 10 (2000) 942–955.

[121] P. Yin, A.M. Tourapis, J. Boyce, Fast mode decision and motion estimation for JVT/H.264, in: Proceedings of IEEE International Conference on Image Processing, Barcelona, Spain, 2003.

[122] Y. Yu, J. Zhou, Y. Wang, C.W. Chen, A novel two-pass VBR coding algorithm for fixed-size storage application, IEEE Trans. Circuits Systems Video Technol. 11 (2001) 345–356.

[123] Z. Zhang, G. Liu, H. Li, Y. Li, A novel PDE-based rate-distortion model for rate control, IEEE Trans. Circuits Systems Video Technol. 15 (2005) 1354–1364.

[124] X.M. Zhang, A. Vetro, Y.Q. Shi, H. Sun, Constant quality constrained rate allocation for FGS-coded video, IEEE Trans. Circuits Systems Video Technol. 13 (2003) 121–130.