Signal Processing: *Image Communication* 29 (2014) 1046–1062

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Window-based rate control for video quality optimization with a novel INTER-dependent rate-distortion model

Yuan Li [a], Huizhu Jia [a,*], Chuang Zhu [a], Mingyuan Yang [b], Xiaodong Xie [a], Wen Gao [a]

[a] *National Engineering Laboratory for Video Technology, Peking University, Beijing 100871, China*
[b] *Beijing BOYA-HUALU Technology Inc, Beijing 100080, China*

ARTICLE INFO

Article history:
Received 24 March 2014
Received in revised form
8 August 2014
Accepted 16 September 2014
Available online 23 September 2014

Keywords:
Rate control
Rate–distortion (R–D) model
Video coding
Consistent video quality
Sliding window

ABSTRACT

Most model-based rate control schemes use independent rate-distortion (R–D) models at macroblock (MB) level to represent the relationship among bit rate, distortion and encoding complexity. However the correlations between frames (INTER-dependency) are not well considered for distortion, bit allocation and quantization parameter (QP) decision. In this paper, a novel INTER-dependent R–D model is proposed based on the theoretical analysis of the relationship between the predicted residual of one frame and the distortion of its reference frame. To achieve both bit rate accuracy and consistent video quality, a window-based rate control scheme with two sliding windows is introduced. One window is to group certain previously encoded frames and current frame to control the bit rate and buffer delay; the other is to group certain future encoding frames to optimize the fluctuation of video quality. Furthermore, the optimization of Lagrange multiplier is also discussed under the INTER-dependent situation. Experimental results demonstrate that the proposed window-based rate control scheme with INTER-dependent R–D model can achieve accurate target bit rate and improve PSNR performance, meanwhile the variation of PSNR is the smallest compared with other three benchmark algorithms. This one-pass rate control scheme is highly practical for the real-time video coding applications.

1. Introduction

Rate control is essential for the real applications of the modern video coding standards such as MPEG-2 [1], H.264/AVC [2] and AVS [3]. Various video coding applications introduce strict bit rate constraint to the bit stream due to the limited transmission bandwidth or storage size. Rate control scheme is responsible for achieving the bit budget by adjusting the quantization parameters (QPs) or trading the compressed video quality in a video encoder. Beside rate, other optimizations are also needed to be considered such as system latency, buffer occupation and smoothness in objective or subjective video quality.

To achieve the bit rate constraint, two essential steps are adopted in a typical rate control scheme, which are frame bit allocation and QP decision. The former is used to allocate bit quota among different video frames, which usually takes the buffer latency, video quality and coding complexity into consideration. The latter, after allocation of the frame bit budget, decides the QP of a frame or each MB of the frame to achieve the bit budget accurately. R–D model is widely applied in this step to represent the relationship among bit rate, distortion and QP.

Many rate control schemes towards these two problems are proposed and developed in the literature. For the frame

* Corresponding author.

http://dx.doi.org/10.1016/j.image.2014.09.004
0923-5965/

bit allocation, the state-of-the-art works can be classified into three categories. The first category allocates equal or nearly equal bits among different video frames. In [4], Ribas-Corbera and Lei used a nearly constant frame bit target to ach;ieve the low buffer delay, meanwhile avoiding the underflow of the buffer. He [5] et al. adopted the similar near-constant frame bit allocation and smoothed the rate variation by adjusting the distortion in a small range. These methods can maintain the small fluctuation of the encoder buffer. However, the source video complexity is not well considered in these schemes, which will cause the fluctuation of the video quality especially when high motion occurs or scene changes. The second category assumes that the video content is stationary among different group of pictures (GOP). Under this assumption, equal bits are allocated to each GOP. Within a GOP, the frame bit allocation scheme assigns a fixed weighting factor according to the frame type. The typical one of this category is TM5 [6] for MPEG-2, it also considers the fullness of the compressed bit buffer. The JM [7] for H.264/AVC uses the similar mechanism for frame bit allocation, meanwhile taking the hypothetical reference decoder (HRD) [8] buffer into consideration to further regulate the frame bit budget. Since the assumption of these methods is not always true, the fluctuation of the video quality is unavoidable. The third category is aimed to achieve smooth video quality. The basic idea of these methods is to allocate more bits to the high-complexity frames and less bits to the low-complexity frames. In [9], Xie and Zeng proposed a sequence-based bit allocation scheme by tracking the non-stationary characteristics in a video sequence. Xu [10] et al. proposed a window model about the picture quality and the buffer occupancy. By applying window-level bit allocation, the tradeoff between quality smoothness and buffer smoothness can be achieved. The "*forward*" rate control scheme, which means that allocating frame bits based on the characteristics of the current frame or future frames via certain pre-analysis, is widely used in these methods. Our work, towards the one-pass real-time encoding application with smooth video quality, also belongs to this category.

For QP decision, various R–D models are proposed in the literature. Some of these R–D models [4,5,11–14] assume that the video coding units are independent with each other. Under this assumption, Chiang and Zhang [11] proposed a quadratic R–D model to calculate the target bit rate for each frame, which was adopted in both MPEG-4 and H.264/AVC. In [4], a MB-level R–D model was used to choose the QP, meanwhile the Lagrange optimization was introduced to minimize distortion. The rate control scheme was adopted by H.263. He et al. [5] proposed a linear ρ-domain R–D model, which used the percentage of zero coefficients after quantization to approximate the bit rate. To tackle the inherent dilemma between rate control and R–D optimization (RDO) in H.264/AVC, Ma et al. [12] used the true quantization step size to establish the R–D model and proposed a rate control scheme with partial two-pass process at MB level [13] proposed an enhanced R–D model, which modeled the source bits as the function of the quantization step size and the complexity of coded 4×4 blocks. In [14], a linear model was formulated to describe the relationship between the total amount of bits for both texture and non-texture information and the QP. There are also a number of investigations for R–D models in the next generation video coding standard—High Efficiency Video Coding (HEVC). Based on the quadratic R–D model in [11], Choi et al. [27] proposed a pixel-based unified rate-quantization (URQ) model, which employed a mean of absolute difference (MAD) factor to predict the texture complexity. This rate control algorithm was adopted in the HEVC test model reference software version 6.0 (HM6.0) [28]. In [29], a rate control scheme using a linear R-λ model was proposed, which showed smaller bit rate errors than the URQ model and was adopted in HM10.0 [30]. Considering the quadtree coding structure in HEVC, Seo et al. [31] proposed a rate control scheme with a new R–D model based on the Laplacian function to minimize the fluctuation of video quality. In [32], a frame-level rate control scheme based on texture and nontexture rate models was proposed, which considered the different statistical characteristics of transform coefficients depending on the depth levels of coding units (CUs). A better R–D performance could also be achieved compared to the previous methods. However, in the more general case, the coding units may not be coded independently, especially when the INTER-dependency is taken into consideration. Here INTER-dependency means that both distortion and bit rate of the current encoding inter frame (either P or B frame) are highly correlated with the distortion of its reference frame, because of the prediction process between the inter frame and its reference frame. To tackle the rate control problem with INTER-dependent characteristics, Ramchandran et al. [15] provided a trellis-based solution for an arbitrary set of QPs for each coding unit. The computational complexity grew exponentially with the increase of dependent frame numbers. In [16], Lin and Ortega used interpolation to establish the approximated R–D curves. The spline interpolation and piecewise linear interpolation were adopted for I frames and P frames respectively. Liu et al. [17] analyzed the dependent temporal-spatial bit allocation problem and proposed two iteration algorithms to reduce the computational complexity. In scalable video coding, Liu and Kuo [18] proposed a GOP-based distortion model for different temporal layers according to the dependency between the base layer and the enhancement layer. The algorithms of [16–18] need to encode the source video several times, which are not suitable for real-time applications.

The above-mentioned R–D models are established by heuristic analyses and statistical examinations. However, the theoretical INTER-dependent R–D model among different coding units needs to be further developed. In this paper, we first analyze the INTER-dependent problem and establish the relationship between the residual of one frame and the distortion of its reference frame. Based on this analysis, we derive the INTER-dependent distortion-quantization (D–Q) model and rate-quantization (R–Q) model via the study of the spatial-domain residual and the transform-domain residual. Then a window-based rate control scheme is proposed with the complexity-based frame bit allocation and video quality optimization. Furthermore, the optimization of Lagrange multiplier is also discussed under the INTER-dependent situation. Experimental results demonstrate that the proposed window-based rate control scheme with INTER-dep-