Signal Processing: *Image Communication* 29 (2014) 1029–1045

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Adaptive bit allocation scheme for rate control in high efficiency video coding with initial quantization parameter determination

Heng Sun*, Shaoshuai Gao, Can Zhang

School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form
16 September 2014
Accepted 17 September 2014
Available online 28 September 2014

Keywords:
HEVC
Rate control
Texture complexity
Bit allocation
Quantization parameter clip
Adaptive initial quantization parameter

ABSTRACT

One of the challenges in rate control (RC) lies in how to efficiently determine a target bit rate that will be used for the quantization parameter (Q_p) calculation process during video coding. In this paper, we investigate the issues over the existing bit allocation algorithms for the RC process in High Efficiency Video Coding (HEVC) and propose an complexity-based bit allocation scheme to improve the encoding performance. First, we model the relationship between encoding bit rate and texture complexity by a linear rate function. Second, compared with traditional complexity estimation methods, a more accurate model is proposed to measure the texture complexity considering the spatial–temporal correlations. Third, based on the proposed rate function and texture complexity measurement model, we develop an adaptive bit allocation scheme for RC in HEVC. At the same time, depending on the encoder buffer status, an adaptive Q_p clip range determination algorithm is also developed to achieve the encoding quality smoothness while keeping the bit rate fluctuation at an acceptable level. Then, we exploit to determine the initial Q_p efficiently and adaptively according to video contents. Experimental results demonstrate that the proposed RC algorithm can achieve better rate-distortion (R–D) and rate-control performance than that of the state-of-the-art RC scheme implemented in the HEVC reference software HM11.0.

1. Introduction

Rate Control (RC) plays a key role for ensuring effective channel adaption during video delivery. It aims to achieve good visual quality under a certain transmission bandwidth constraint. In general, a typical RC scheme consists of two basic operations: (1) bit allocation, i.e., how to properly allocate target bits to each basic unit, and (2) bit-rate control, i.e., how to adjust the quantization parameters to properly encode each basic unit to achieve the allocated bits. For video coding, the basic unit is the basis based on which RC is resolved and for which distinct values of quantization parameter (Q_p) are calculated.

For bit allocation operation, the main problem is how to efficiently distribute the bits budget among image blocks to achieve the best rate–distortion (R–D) performance. Currently, the bit allocation schemes can be roughly classified into two categories. The first is predominantly based on buffer status and the second uses a complexity-based approach. Many RC schemes, such as MPEG-4 Q2 [1–3], the linear RC scheme [4] and TMN8 [5], belong to the first category. For instance, in [4], the target bit rate for each frame is determined by $R/f-b$, where R, given to encoding sequence in bits per second (bps), is the target bit rate determined by current available channel bandwidth, f is the predefined frame rate in frames per second (fps), and b is the value determined by current coded picture buffer

---

* Corresponding author.
E-mail addresses: hengsunheng@gmail.com (H. Sun), ssgao@ucas.ac.cn (S. Gao), czhang@ucas.ac.cn (C. Zhang).

http://dx.doi.org/10.1016/j.image.2014.09.006
0923-5965/

status. Meanwhile, in order to guarantee the coded picture buffer of a video encoder neither overflows nor underflows, the final allocated bits for each frame should be clipped with an upper and lower bounds which are both determined by the coded picture buffer status. When the buffer level of the coded picture buffer exceeds a predetermined buffer level, frames are forced to skip to keep the buffer from an overflow. However, because the video texture complexity of a video sequence may vary from frame to frame due to the changes in video activity, video encoding quality will vary dramatically if only buffer-based bit allocation scheme is used. As a result, such bit allocation schemes are unable to achieve good R–D performance because they do not match the time-varying characteristics of video signals. In [6,7], the target bits are allocated among video frames by a multi-pass encoding process. Nevertheless, due to the time-consuming multi-pass procedure, it is more appropriate for off-line applications. Other RC algorithms, such as [8–10], perform bit allocation according to the video texture complexity. Its basic idea is that each basic unit with higher texture complexity can be allocated more bits; whereas, each basic unit with lower texture complexity may be allocated fewer bits. For example, in [9], an optimized method was proposed to assign target bits to each macroblock (MB) according to its mean absolute difference (MAD) statistics predicted by using the actual MAD of collocated MB in the previous coded frame as the same type of current frame. In such a case, are only temporal correlations used. Unfortunately, such texture complexity prediction schemes usually suffer from relative large prediction errors in case of low temporal correlations available. Therefore, such bit allocation schemes also fail to achieve good R–D performance due to the inaccurate video texture complexity estimation.

Traditionally, for bit-rate control operation, RC regulates the coded video bit stream mainly by adjusting Q_p. To achieve the target bit rate, rate-quantization (R–Q) models are often employed for representing the encoding bit rate by means of Q_p and other parameters, such as the MAD of a residual MB [3,9], and the percentage of zero quantized transform coefficients [4]. However, using parameters such as MAD for R–Q modeling causes the chicken-and-egg dilemma [11] to high quality video coding standards because the Lagrangian coder control method incorporated into video encoders demands Q_p be evaluated before intra/interprediction, but until the end of intra/interprediction, RC cannot access the statics such as MAD which is required for the Q_p calculation. This interdependency between rate-distortion optimization (RDO) and RC makes RC more challenging than previous standards. To overcome the chicken-and-egg dilemma, Ma et al. [3] propose a partial two-pass RC scheme with a linear R–Q model being proposed. In [9], the authors propose a linear distortion–quantization (D–Q) model with a linear R–Q model and hence develop a R–D joint optimized solution to Q_p determination.

The High Efficiency Video Coding (HEVC) standard [12] is the latest video coding standard which can significantly improve the coding efficiency over its preceding standards such as MPEG-2, H.263, MPEG-4 and H.264/AVC [13] because it has incorporated many new coding tools which have been not included in the conventional standard video codecs. The core of the coding layer in HEVC is the coding tree unit (CTU) structure which has a size selected by video encoder and can be larger than a traditional MB which is the analogous structure in previous standards [12]. CTU is also referred to as the largest coding unit (LCU) whose size can be between 16 × 16 and 64 × 64 pixels with a larger size usually increasing coding efficiency.

1.1. Rate control in HEVC

Similar to the prior video standards, HEVC also recommends its own RC schemes. Based on the new features of HEVC and the R–Q model proposed in [14], a pixel-wise unified R–Q model is proposed in [15] named quadratic pixel-wise unified R–Q (URQ) model, which is the recommended RC scheme in the HEVC reference software HM6.1. It calculates Q_p based on the predicted target bits and image complexity before actual encoding. In [16,17], the improvement of the pixel-wise URQ RC performance is achieved by using quantization step size instead of Q_p value. In [18], a RC scheme based on a linear $R-\lambda$ model is proposed, which shows less bit-rate mismatching and better R–D performance than the RC scheme based on the pixel-wise URQ model. In [19], the $R-\lambda$ model based RC scheme was further extended into an intra-frame RC scheme based on sum of absolute transformed difference (SATD) to achieve more accurate matching of target bit rate for intraframes. In the state-of-the-art RC scheme for HEVC, the $R-\lambda$ model based RC scheme [18] is used for inter-frame coding and the RC scheme in [19] is adopted for intra-frame coding, which is implemented in the HEVC reference software HM11.0. In [17], MAD is used as complexity measure and determined by a temporal linear prediction formula. In [18], the $R-\lambda$ model based RC scheme also employs MAD, which is predicted by that in the collocated position of the previous coded frame belonging to the same level of current frame, to characterize the complexity of residual signals. Although this texture complexity estimation method is simple, yet it is not accurate enough. As mentioned in the previous subsections, the prediction performance of such texture complexity estimation schemes is weakened by the poor use of spatial–temporal correlations. A better texture complexity estimation scheme should be adaptive to not only temporal contents but also spatial contents. As a result, using MAD to estimate the texture complexity of video sequences could not adapt well to the changing contents of video sequences. In [19], SATD is used to measure the complexity of intra-frame. In the state-of-the-art RC scheme implemented in HM11.0, it allocates target bits to every frame according to a predefined proportion. However, as previously stated, such bit allocation schemes cannot adapt to different videos dynamically. The complexity of each frame has to be considered for good coding performance. Therefore, before performing the bit allocation operation, texture complexity measurement model should be proposed first. Unfortunately, there are few works about how to accurately measure video texture complexity according to the properties of input video sequences.

The RC schemes mentioned above mainly focus on how to employ R–Q or D–Q models for improving the RC performance of HEVC; whereas, another important factor influencing the RC performance, i.e., how to determine the initial Q_p value, has not been well addressed yet. In the traditional RC schemes, the initial Q_p value is determined