J. Vis. Commun. Image R. 35 (2016) 120–131

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Low complexity encoder optimization for HEVC ☆

Shanshe Wang [a,*], Falei Luo [c], Siwei Ma [a], Xiang Zhang [a], Shiqi Wang [a], Debin Zhao [b], Wen Gao [a]

[a] Institute of Digital Media & Cooperative Medianet Innovation Center, Peking University, Beijing, China
[b] Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[c] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Article history:
Received 30 July 2015
Accepted 9 December 2015
Available online 17 December 2015

Keywords:
HEVC
Low complexity
Intra mode decision
CU splitting
Reference frame selection
Rate distortion optimization
Video coding
Encoder optimization

ABSTRACT

High Efficiency Video Coding (HEVC) improved the coding efficiency significantly. Compared to its predecessor H.264/AVC, it can provide equivalent subjective quality with more than 57% bit rate reduction. However, the improvement on coding efficiency is obtained at the expense of much more intensive computation complexity. In this paper, based on an overall analysis of computation complexity at the HEVC encoder, a low complexity encoder optimization scheme is proposed by reducing the number of available candidates for evaluation in terms of the intra prediction mode decision, early termination of coding unit (CU) splitting and adaptive reference frame selection. With the proposed scheme, the rate distortion optimization (RDO) technique of HEVC can be implemented in a low-complexity way for complexity-constrained encoders. Experimental results demonstrate that, compared with the original HEVC reference encoder implementation, the proposed optimization scheme can achieve more than 40% complexity reduction on average with coding performance degradation as only 0.43% which can be ignorable.

1. Introduction

The dramatically increasing of high definition (HD) and beyond-HD (e.g. 4k × 2 k or 8k × 4 k) videos are creating stronger demand of high efficiency video coding technology, which are beyond the capabilities of the state-of-the-art video coding standard such as H.264/AVC [1]. Therefore, the ITU-T Video Coding Expert Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) formed the Joint Collaborative Team on Video Coding (JCT-VC) in April 2010 to develop the new coding standard which is formally published in 2013 with the name as High Efficiency Video Coding (HEVC) [2] or H.265.

HEVC shows great improvement in many aspects. In [32], the authors provide an overview over the new characteristics which are likely to be used in HEVC in wireless environments and discusses several research challenges. Compared with the similar applications in H.264/264 [33], HEVC codec is more effective in both subjective and objective visual quality. In [31], the authors detail the HEVC applications in video transport and delivery such as broadcast, television over the Internet Protocol, Internet streaming, video conversation, and storage. It can be concluded that HEVC provides more flexible coding structure based on different system layer than those in H.264/AVC [34].

For the coding performance improvement, many new coding tools or coding structures are adopted or improved, which bring great performance improvement. Compared to its predecessor H.264/AVC, HEVC achieved more than 57% bit rate saving in terms of perceptual quality [3,4]. However, the improvement on coding efficiency is obtained at the expense of more intensive computation complexity. In [8,9], the authors analyzed the implementation complexity and the coding efficiency of these advanced coding tools in HEVC.

For evaluating the compression efficiency of each candidate configuration, the encoder usually employs the Lagrange multiplier optimization technique [6], which is expressed by

$$\min\{J\}\ J = D + \lambda \cdot R, \tag{1}$$

where J is the Lagrange rate–distortion (R–D) cost function to be minimized, D and R are the reconstruction distortion and entropy coding bits of a certain unit and λ is the Lagrange multiplier. The minimization process of the R–D cost is the well-known rate–distortion optimization (RDO). In general, to obtain accurate D and R, for each candidate, the encoder has to perform transform, quantization, entropy coding, inverse quantization, inverse transform, and pixel reconstruction, which makes the R–D cost calculation very time-consuming and brings great burden to the encoder implementation [7]. However, with the rapid developments of the portable devices, the discrepancy between the computationally intensive video codec and the limited computational capability of

☆ This paper has been recommended for acceptance by M.T. Sun.
* Corresponding author.

http://dx.doi.org/10.1016/j.jvcir.2015.12.005
1047-3203/

hardware platforms has become a bottleneck, especially in the real time visual communications. In view of these requirements, some researches on low complexity implementation of HEVC encoder have been done based on the new adopted coding tools or new coding structures.

For intra coding optimization of HEVC, many fast algorithms are proposed to reduce the coding complexity of intra mode decision. In [17], a partition depth prediction scheme is proposed based on the spatial correlation among adjacent prediction unit (PU). An adaptive threshold updating method is provided in order to improve the prediction accuracy. Due to the fact that the gradient can represent some characteristics of video content, in [18], a gradient based optimization scheme for intra mode decision is proposed. In [19], according to the correlation between the texture of video and the optimal mode, a fast intra mode decision method is presented. In [20], the authors proposed a fast intra mode decision for HEVC, in which micro level and macro level decision algorithms were presented. At the micro level, Hadamard transform based R–D cost is utilized to reduce the candidate number and an early rate distortion optimization quantization (RDOQ) skip method is also introduced to further reduce the coding complexity. At the macro level, an early coding unit (CU) split termination method is provided in terms of estimated R–D cost. In [21], a fast intra decision method is proposed to reduce the number of intra-prediction modes in rough mode decision (RMD) process based on direction information of the co-located neighboring block of previous frame along with neighboring blocks of current frame. In [22], according to analyze the distribution and changing trend of the costs generated by RMD, a fast intra decision scheme is presented to reduce the number of the candidates for the RDO process. In [17], Shen et al. proposed a fast scheme for intra coding in HEVC in order to determine the size of CU based on the depth of the surrounding CUs. Furthermore, a mode decision method is also provided for intra coding. in [30], variance based optimization scheme is incorporated for the intra coding optimization.

In HEVC, the most time consuming component is inter coding. For inter coding optimization, the researches mainly focus on the motion estimation and the quad-tree coding structure including CU, PU and TU. Many optimization schemes have been proposed. In our previous work [23], we propose an adaptive reference frame selection scheme to accelerate the motion estimation process. Since the CU splitting increases much complexity on HEVC, many optimization algorithms are proposed for the early termination of CU and determination of CU depth. In [25], the authors present a fast algorithm for early CU splitting and pruning according to a Bayes decision rule method. In [26], using the available side information such as sample adaptive offset (SAO) parameter values, PU sizes and coded block flag (cbf) data, a fast decision method of CU partition is proposed in order to reduce the coding complexity. In [24], based on the spatial and temporal correlation between the current CU and its adjacent and co-located CU, the authors proposed a fast algorithm which consists of splitting decision and termination decision for CU. For the early termination of CU splitting, a fast algorithm is proposed in [27] based on the energy of prediction residuals. In [11], an early TU decision algorithm for high efficiency video coding is proposed. In [12], coded block flag (cbf) is used to terminate PU encoding process. If the cbf of an inter PU in a CU is zero for luma and chroma except for inter NxN PU, the next PU encoding process for the current CU will be terminated. In [13], another scheme is proposed for skip mode early termination with the aim of optimizing the inter coding process. The basic idea is that if skip mode is the locally optimal mode of the current CU depth, skip mode is then considered to be global optimal mode and sub-tree computation process can be skipped. In [10], the authors proposed a complexity control algorithm for HEVC by adaptively choosing the CU splitting depth.

Though these techniques have shown good coding efficiency, the whole structure of the encoder has not been fully considered. In our previous work [29], we propose a low complexity rate distortion optimization scheme combining intra optimization and inter optimization to reduce the coding complexity of HEVC. In this paper, based on an overall analysis of computation complexity in HEVC encoder, we further improved the low complexity encoder optimization scheme for LDP (Low delay P) configuration of HEVC. Compared to our previous work, the main contributions mainly includes: in intra prediction, the gradient variance related to the video content is utilized to the speed up intra mode decision; in inter CU decision, the correlation between the rate distortion cost and CU splitting are exploited to accelerate the CU splitting termination process; and for reference frame selection, a probability model is proposed based on the spatial and temporal correlations among the adjacent frames and CUs, then the model is employed to shrink the RFS (Reference frame set) in order to reduce the inter coding complexity. The architecture of our scheme is shown in Fig. 1. It can be observed that our algorithms apply to the modules that require intensive computation in the encoder.

The rest of this paper is organized as follows. In Section 2, we first give an overall analysis of the computation complexity in HEVC encoder. Then the most time consuming modules in the HEVC encoder are analyzed in terms of the computational complexity and coding efficiency in details, including the block partitioning structure as well as inter and intra prediction in HEVC. In Section 3, we propose a low complexity RDO scheme for intra mode decision. In Section 4, a low complexity RDO scheme for inter coding is proposed including fast reference frame selection and inter CU splitting decision. Experimental results are provided in Section 5, and Section 6 concludes the paper.

2. Complexity analysis of HEVC encoder

2.1. Encoder computational complexity analysis

HEVC is based on traditional hybrid prediction/transform coding framework as described in Fig. 2 [3].

The newly adopted coding structure in HEVC offers more possibilities to split a frame into multiple units and more ways of combining different coding tools and parameters. Though this doesn't have significant impact on the decoder from the complexity aspect, it imposes a heavy computation burden to the encoder by fully leveraging its capabilities. Experiments are conducted to show the time consumption of the major modules in the encoder for Random Access and Low Delay configurations as illustrated in Table 1. It can be seen that the most time-consuming part is the motion estimation as a result of the multiple reference motion compensation.

For all-intra configuration, the coding complexity mainly comes from the mode decision from all available candidate modes. It can be seen from Table 2 that in the original configuration of HEVC, the percentage of intra coding complexity is over 80%. It can be also observed that when the number of intra prediction mode is reduced, the coding complexity can also be reduced much. E.g. when the number of mode candidates for PU with different size changes from the original S to different sets: S1, S2 and S3 as shown in Table 2, the coding complexity can be reduced by 8.50%, 16.0% and 25.7% respectively. Thus the number of candidates has crucial influence for the coding complexity. Moreover, it is reported that the full RD search of all available candidate modes will only achieve −0.4% coding gain, but bring about three times of the encoding time of the fast mode decision algorithm [14].

Based on the above considerations, in this section, the relationship between the compression efficiency and the corresponding