J. Vis. Commun. Image R. 40 (2016) 34–41

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Computational complexity allocation and control for inter-coding of high efficiency video coding with fast coding unit split decision ☆

Jiunn-Tsair Fang [a], Zong-Yi Chen [b], Chang-Rui Lai [b], Pao-Chi Chang [b,*]

[a] Ming Chuan University, Department of Electronic Engineering, No. 5, Deming Rd., Taoyuan City 33348, Taiwan
[b] National Central University, Department of Communication Engineering, No. 300, Jhongda Rd., Taoyuan City 32001, Taiwan

ARTICLE INFO

Article history:
Received 19 January 2016
Revised 7 June 2016
Accepted 13 June 2016
Available online 15 June 2016

Keywords:
High-efficiency video coding (HEVC)
Complexity allocation
Complexity control
Coding unit
Motion vector

ABSTRACT

HEVC provides the quadtree structure of the coding unit (CU) with four coding-tree depths to facilitate high coding efficiency. However, compared with previous standards, the HEVC encoder increases computational complexity considerably, thus making it inappropriate for applications in power-constrained devices. This study therefore proposes a computational complexity allocation and control method for the low-delay P-frame configuration of the HEVC encoder. The complexity allocation includes the group of pictures (GOP) layer, the frame layer, and the CU layer in the HEVC encoder. Each layer involved uses individual method to distribute the complexity. In particular, motion vector estimation information is applied for CU complexity allocation and depth split determination. The total computational complexity can thus be reduced to 80% and 60% or even lower. Experiment results revealed that the average BD-PSNR exhibited a decrease of approximately 0.1 dB and a BD-bitrate increment of 2% when the target complexity was reduced to 60%.

1. Introduction

In recent years, mobile phones have become popular consumer electronics. Smartphones are ubiquitous in daily life for accessing the Internet for videos or multimedia applications. People are increasingly capturing and sharing multimedia on mobile devices; therefore, the demand for high-quality and real-time video is increasing. Hence, complex algorithms with heavy computations have been developed in recent video standards. Heavy computations, however, consume considerable power and reduce battery life, hindering video applications in power-constrained mobile devices. Therefore, reducing the computational complexity is crucial for video applications in mobile devices.

The video standard H.264/AVC employs the macroblock (MB) as the largest block for prediction coding [1]. An MB is a 16×16 block of pixels, and its prediction modes, partitioned from an MB with different block sizes, are used for motion estimation (ME). The optimal prediction mode is determined according to the rate-distortion optimisation (RDO) procedure. RDO reduces the prediction error but increases the computational complexity. Accordingly, mode decision (MD) or ME are popular topics for complexity reduction in H.264/AVC [2–5].

The newest video standard high-efficiency video coding, (HEVC), finalised in 2013, supports a coding efficiency that is higher than those of previous standards, especially for high-resolution video content [6]. HEVC uses quadtree structures of coding units (CUs) and four coding-tree depths to facilitate high coding efficiency. The largest coding unit (LCU) is composed of a block of 64×64 pixels, which is 16 times greater than the MB block size. The computational complexity is increased considerably, and many fast algorithms have been proposed based on CU depth decision in recent years [7–16].

Although fast algorithms for CU depth decision are intended to reduce the complexity of the HEVC encoder, the main objectives are to achieve a trade-off between the transmission quality and complexity. By contrast, the complexity control method is used for effectively maintaining the transmission quality as the target complexity is downscaled. The purpose of complexity control is not only to reduce the power consumed for a given target complexity but also to effectively control the transmission quality of a power-constrained device.

Extensive studies have been proposed to reduce complexity for H.264/AVC and HEVC, respectively. By contract, studies on complexity control were not as much. We list some representative work as follows. He et al. proposed a theoretical model for

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

* Corresponding author.
E-mail address: pcchang@ce.ncu.edu.tw (P.-C. Chang).

http://dx.doi.org/10.1016/j.jvcir.2016.06.004
1047-3203/

power-RD analysis. They adjusted complexity control parameters to match the available energy supply while maximizing the picture quality [17]. Chien et al. measured the coding gains of coding tools to reorder the steps of the encoding process. Some coding tools with less coding gains may be skipped to meet the required complexity constraint [18]. Corrêa et al. separated the sequence into constrained and unconstrained frames. Information obtained from the collocated areas in the previously encoded unconstrained frames could be used to predict the number of constrained frames and their coding-tree depths [19]. Several studies have investigated the improvement of the prediction performance [20–23]. Moreover, a workload management strategy was proposed for maintaining the frame rate transmission [24]. In [25], Zhang et al. proposed the motion collision count (MCC) to estimate the number of prediction units (PUs) located in the current encoding CU, and then determined the number of CU splits from each CU depth. The complexity control was based on the percentage of CU splits. Recently, a subjective driven complexity control was proposed to reduce and control the encoding complexity of HEVC [26].

This study focuses on the low-delay P-frame (LDP) configuration of HEVC, because this configuration is suitable for applications in low-power devices. The complexity allocation includes the group of pictures (GOP) layer, the frame layer, and the CU layer of the HEVC encoder. Instead of allocating exact computational complexity among frames or CUs, their relative complexity of consumption ratio was estimated. Different methods were applied to different coding layers to estimate the relative complexity consumption. As the target complexity was set, the relative complexity allocation for each coding layer could be preserved, and the complexity control could thus be applied effectively. To further increase the coding efficiency, a formula based on motion vector estimation was used for the fast decision of the CU depth split. Therefore, the proposed method could not only maintain the transmission quality adaptively but also and increase the coding efficiency of HEVC.

The remainder of this paper is organized as follows. Section 2 describes the proposed complexity control method, and the experimental results of are presented in Section 3. Finally, Section 4 presents the conclusion.

2. Complexity control method

In this study, the computational complexity was allocated to each coding layer of HEVC for complexity control. Methods for allocation are detailed in the following subsections. The computational complexity was measured by the operation of the central processing unit (CPU) clock cycles. The time consumption equals the consumed number clock cycles divided by the CPU frequency. First, the complexity was equally distributed among GOPs. Second, each frame complexity was estimated using a parabolic curve at different quantisation parameter (QP) settings. Third, the complexity distribution to each CU depended on the estimated number of PUs located in the current encoding CU block. Fourth, a formula was used for the fast decision of the LCU depth splitting. Finally, the computational complexity was compensated: Any over- or underallocated complexities were equally distributed among the remaining uncoded CUs or frames.

2.1. GOP layer complexity allocation

This study focused on the LDP configuration in the HEVC encoder [27]. The LDP configuration consisted of one I-frame, with the remainder being P-frames. Fig. 1 illustrates the LDP configuration, where each GOP contains four P-frames. The QP settings for the first and third frames were QP + 3, for the second frame, QP + 2,

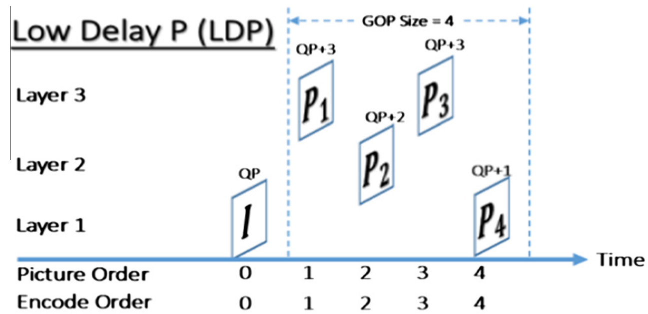

Fig. 1. Hierarchical coding structure of the LDP configuration [13].

and for the fourth frame, QP + 1. This QP setting was the same for the remaining GOPs. Fig. 1 also shows that the encoder order is the same as the picture order, meaning that this configuration was low-delay for real-time transmission [13].

To estimate the computational complexity in the GOP layer, four test sequences from different classes were simulated. The simulation environment is presented in Table 1. The reference software for HEVC was HM 12.1, and eight QP settings were selected. The computational complexity was determined according to the CPU clock. The sequences of Class D were excluded in this study because of their low resolution.

Fig. 2 shows the complexity consumption of the first 97 frames with QP 27 for four sequences (Classes A, B, C, and E). The complexity consumption depended on the frame size and sequence content. Large frame sizes with high-motion sequences, such as Class A_PeopleOnStreet, consumed more complexity than small frame sizes with low-motion sequences, such as Class E_Vidyol. Fig. 2 also shows a crucial characteristic of the LDP configuration: The complexity consumption for each sequence could be treated as a periodic signal with a period of 4, except for the first GOP. In other words, processing each GOP, composed of four frames, involved almost the same complexity.

Note that processing the first I frame involves a lower complexity compared with processing other frames. This is because P frame uses interprediction and I frame uses intraprediction for prediction coding. It usually takes more complexity to process interprediction than to process intraprediciton. From HEVC default setting, four reference frames are used for interprediction. The first four P frames (the first GOP) thus use less number of reference frames for interprediction, and involves a lower complexity compared with processing the other GOPs. For these four P frames, their complexity consumption increases as the number of available reference frames increases.

According to the LDP configuration, the complexity of all P frames, C_P, can be expressed as follows:

$$C_P = C_E - C_I \quad (1)$$

Table 1
Simulation environment.

Reference software	HM 12.1
Sequence	Class A_PeopleOnStreet
	Class B_Kimono1
	Class C_RaceHorse
	Class E_Vidyo1
FramesToBeEncoded	97
Configuration file	Low Delay P (IPPP...)
QP	20, 22, 25, 27, 30, 32, 35, 37, 40, 42
Hardware	Intel(R) Core(TM) i7 CPU 920 @ 2.67 GHz,4.0 GB of RAM