Signal Processing: Image Communication 62 (2018) 74–81

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Efficient screen content intra coding based on statistical learning

Hao Yang [a], Liquan Shen [b,*], Ping An [b]

[a] Key Laboratory of Specialty Fiber Optics and Optical Access Networks, School of Communication and Information Engineering, Shanghai University, China

[b] Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, China

ARTICLE INFO

Keywords:
Screen content coding
Intra prediction
Fast algorithm
Statistical learning

ABSTRACT

Screen content videos play an important role in recent popular mobile technologies and cloud applications. With the new coding tools (including intra block copy, palette mode etc.) adapted, high efficiency video coding (HEVC) based screen content coding (SCC) achieves high coding efficiency but requires a very high computational complexity. In this paper, we propose a fast intra coding algorithm for SCC based on statistical learning. First, we analyze the complexity distribution of SCC encoders. Then, a classifier is designed to determine whether current coding unit (CU) should be split into four sub-CUs without performing intra prediction procedure at current CU depth level or not. If current CU should not be split into sub-CUs, another classifier is used to decide either 35 traditional intra modes or SCC modes (i.e. intra block copy mode and palette mode) will be performed at current CU depth level. Three kinds of features are chosen to generate the classifiers: (1) coding information of current CU; (2) texture information within current CU; (3) information of nearby CUs. Experimental results show that the proposed algorithm achieves a computational complexity reduction with 45% on average, and only 1.6% BDBR increase compared to the original coding in SCC reference software.

1. Introduction

High efficiency video coding has been developed for years by Joint Collaborative Team on Video Coding (JCT-VC), which organized by ITU-T Video Coding experts Group (VCEG) and ISO/IEC Moving Pictures Expert Group (MPEG). It achieves 50% bit-rate saving for the same perceptual quality relative to the performance of H.264/Advanced Video Coding (AVC) [1]. In recent years, mobile technologies and cloud applications became more and more popular, such as remote desktop, screen sharing applications, wireless display, online learning, etc. These technologies and applications create a new kind of video, which contains a significant amount of stationary or moving computer graphics and text, called screen content videos. To efficiently compress screen content videos, the screen content coding extension is developed based on the latest HEVC since January 2014 [2].

SCC inherits the flexible quad-trees partition structure from HEVC. Besides, three major coding tools have adopted in SCC, known as adaptive color transform (ACT) [3,4], palette mode (PLT) [5] and intra block copy (IBC) [6,7]. These new coding technologies achieve more than 50% bit-rate saving for screen content comparing with HEVC [8], which also put heavy computational burden on SCC encoders. Recently, a number of fast coding algorithms [9–20] have been proposed to reduce the computation complexity of HEVC encoders. They can be roughly classified into two categories: fast mode decision methods and fast CU size decision algorithms.

Fast mode decision methods are proposed in [9–13]. A gradient based fast mode decision algorithm is proposed in [9], which calculates the gradient directions prior to intra prediction and a gradient orientation histogram is generated for each CU. Based on the distribution of histogram, only a small part of candidates are chosen for rate-distortion optimization (RDO). The rough modes decision (RMD) strategy in [10] reduces the complexity by selecting several most promising candidate modes using the sum of absolute Hadamard transform difference (SATD). The fast intra prediction scheme in [11] arranges the candidate modes into different groups based on the average gradient in horizontal direction and vertical direction, and the homogeneity is early evaluated to detect the unidirectional CUs. The fast intra prediction mode decision strategy in [12] is based on two observations: (1) the RMD costs listed by mode number generally follow the same trend with the RDO costs; (2) the local salient modes, whose RMD costs have a significant drop compared with adjacent modes, tend to be promising competitors for the optimal mode. An early skip mode decision in [13] utilizes the distortion characteristics of 2N×2N Merge mode, and those prediction units, which are decided as the SKIP mode after calculating rate-distortion

* Corresponding author.
E-mail address: jsslq@shu.edu.cn (L. Shen).

https://doi.org/10.1016/j.image.2017.12.001
Received 2 June 2017; Received in revised form 16 October 2017; Accepted 1 December 2017
Available online 14 December 2017
0923-5965/

cost of 2N×2N Merge mode, do not undergo remaining mode decision processes.

Fast CU size decision algorithms are proposed in [14–20]. Dynamic depth range and early termination schemes for fast intra prediction in [14,15] skip some prediction modes which are rarely used in the parent CUs in the upper depth levels or spatially nearby CUs. The fast CU splitting and pruning method in [16] is performed at each CU depth level according to a Bayes decision rule based on low-complexity RD costs and full RD costs. Machine learning methods are also introduced into fast intra coding algorithms in recent years. Learning-based method in [17] uses decision tree (ML-DT) as the data mining classifier. A machine learning based fast CU decision framework is proposed in [18], which employs a three-output joint SVM classifier. The optimal learning parameters are determined by minimizing the coding complexity with given RD cost constraints. Bayesian decision rules is employed in [19] to terminate the CU partitioning process using joint online and offline learning. The fast quad-trees pruning algorithm in [20] uses prediction residuals statistics from current CU to avoid the initial branching of the coding quad-trees.

The aforementioned methods are well developed for HEVC encoders, which achieve significant time reducing. However, they are proposed for natural video coding, which do not consider the characteristics of screen content videos. Different from natural videos, screen content videos frequently contain large flat regions, repeated texture patterns and limited number of colors, which make aforementioned fast coding methods inefficient. Besides, the aforementioned methods do not consider the computational complexity caused by the new SCC coding tools. Therefore, several fast algorithms for screen content coding are explored as follows.

A fast search algorithm in [21] presents three fast intra motion vector search methods based on the regular intra mode cost and CU activity. Hash based fast search in [22] has been adapted to SCC, which finds matching blocks using the hash value at entire coded regions within current frame. CU entropy and coding bits are used in [23] to determine CU splitting decision (CUSD-EC). A fast SCC algorithm in [24] incorporates simple intra prediction approach and redundant bits removal approach (SIP-RBR), which focuses on the frequent smooth regions where the traditional intra modes and IBC mode are skipped according to the CU boundary samples. Machine learning based fast CU partition decision (CUPD-ML) in [25] employs a neural network with CU statistics features and sub-CU homogeneity related features. Fast intra coding algorithm (FMD-FSM) in [26] includes two strategies: (1) fast intra modes prediction (FMD) scheme uses the temporal correlation among CUs; (2) fast search method (FSM) utilizes adaptive search step to further speed up the time-consuming block matching search in IBC. A fast intra prediction method (CPA) in [27] reduces the complexity based on content property analysis, in which the quad-trees partition for homogeneous regions is terminated. When DC or PLANR mode is the best mode after checking traditional intra prediction rough modes, the newly adopted prediction modes (i.e. IBC and PLT) are also skipped.

The most previous SCC works either focus on fast search methods or generally utilize the texture homogeneity to speed up the coding parameters (optimal intra prediction mode and optimal CU depth level) determination process. They do not consider the properties of IBC and PLT which rely on the patterns and colors in the previous coded area. Sometimes the texture-rich regions will be encoded in large CU size using IBC or PLT. Thus, the fixed texture features within the CUs cannot predict the optimal coding parameters precisely, and optimal coding parameters also have high correlation with spatial neighbor CUs and the quantization parameter (QP). To solve this problem, we propose an efficient intra coding algorithm based on three kinds of features, including texture statistical information, coding information of spatial nearby coded CUs and coding information of current CU. In our algorithm, 2N×2N traditional intra mode will be checked first for each CU to get its coding information. Then, two classifiers are learned using above three kinds of features. Each CU will be classified into either the splitting category or the un-splitting category by the first classifier. These CUs, which are classified into the splitting category, will be directly split into four sub-CUs without performing intra prediction at current depth level. The other CUs of un-splitting category will be further classified into screen coded CUs and direction coded CUs by the second classifier. Screen coded CUs skip checking traditional intra prediction modes and only perform SCC coding modes. Direction coded CUs only check 35 traditional intra prediction modes. Experimental results show that the proposed algorithm significantly reduce intra coding time while maintaining coding efficiency.

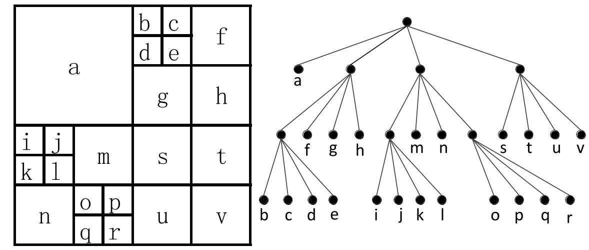

Fig. 1. Hierarchical quad-tree partition structure.

The rest of this paper is organized as follows. Section 2 introduces the intra coding strategy and illustrates coding complexity distribution of screen content coding. Section 3 first analyses the SCC intra modes and CU depth level distributions, and then presents the motivation. Section 4 proposes a machine learning based fast intra coding algorithm, including early CU splitting decision strategy and early intra coding mode detection strategy. Section 5 presents the experimental results and compares performance of proposed algorithm with several state-of-the-art algorithms, and Section 6 concludes the paper.

2. Overview of SCC intra coding

Considering the characteristics of screen content videos, three major coding tools, known as ACT, IBC and PLT, have been adopted to SCC. ACT removes inter-color component redundancy [28], which is only applied for intra directional prediction. Videos in the RGB or YUV color space can be coded directly or adaptively converted to YCoCg color space, and thus causes extra computation complexity. IBC finds the matching block from the reconstructed areas of current picture, and Block Vector between current block and matching block is signaled to decode side. In current HEVC-SCC test software, IBC and INTER modes are unified [29]. IBC is very efficient for coding repeated patterns in text and graphics content, and high computational complexity is also introduced to SCC intra coding. PLT [30] is designed according to the characteristics of screen content which usually contain few color values with large difference. It enumerates these colors by building a color table, and an index is sent to indicate which color each pixel belongs to. PLT is efficient for the blocks with limited color values but bits-consumption for the color-rich regions.

SCC inherits the flexible quad-tree partition structure from HEVC. Input frames are firstly divided into coding tree units (CTU) with block size of 64 × 64 pixels, and CTU can be further recursively partitioned into four equal-size sub-CUs using quad-tree structure until largest CU depth level reaches, as shown in Fig. 1. At each CU depth level of the CU quad-tree, the optimal encoding parameters are derived based on prediction units (PU). Each intra CU has two PU modes: 2N×2N mode and N×N mode. In screen content test model software version 5 (SCM-5.0), the strategy of deriving optimal parameters is shown in Fig. 2.

For each input CTU, traditional intra modes are tested in the first place. Then, the new SCC modes are checked. All candidate modes are evaluated with the cost of rate-distortion optimization (RDO), which is defined as

$$RDCost = D + \lambda \cdot R_{mode} \tag{1}$$