J. Vis. Commun. Image R. 29 (2015) 33–45

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A new intra prediction with adaptive template matching through finite state machine ☆

Chia-Hung Yeh [a], Shu-Jhen Fan Jiang [a], Chih-Yang Lin [b,c,*], Pei-Lun Suei [d], Min-Kuan C. Chang [e]

[a] Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan
[b] Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan
[c] Dept. of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan
[d] Research Center for Information Technology Innovation, Academia Sinica, Taipei 115, Taiwan
[e] Graduate Institute of Communication Engineering, National Chung Hsing University, Taichung 402, Taiwan

ARTICLE INFO

Article history:
Received 15 February 2014
Accepted 18 January 2015
Available online 11 February 2015

Keywords:
Intra prediction
H.264/AVC
High Efficiency Video Coding
Finite state machine
Motion estimation
Video coding
Side-match vector quantization
Codebook

ABSTRACT

This paper presents a new approach that aims to improve the performance of the intra block coding of H.264/AVC and HEVC by using a finite state machine. Based on the high correlations between a frame's neighboring blocks, the finite state machine is employed at both the encoder and decoder to reduce the number of bits required for intra encoding, improving the coding performance of videos. With the matching adaptive template, a better prediction block is found. Through the proposed extra intra prediction modes, the number of bits required to encode a block is reduced significantly, and thus a better intra coding performance is achieved. In addition, an early termination is proposed to speed-up the coding performance. Experimental results show that with the proposed method, the bit rate can be reduced 11% on average when compared to H.264/AVC and 4% on average when compared to HEVC.

1. Introduction

Video coding is one of the most significant fields of multimedia research. In the last two decades, many high quality video coding techniques have been proposed. The High Efficiency Video Coding (HEVC) standard was developed by the Joint Collaborative Team on Video Coding (JCT-VC) in 2010, and was completed in 2013. The HEVC standard is very flexible; it can be used in a wide variety of applications over a diversity of networks and systems [1]. The objective of HEVC is to improve the video quality of previous standards. Some later improvements of HEVC, such as H.264/AVC, introduced new features such as coding tree unit, intra coding, and an in-loop deblocking filter [2–4], to improve the coding performance of videos. Among them, intra coding has one of the most important coding functions, as it prevents error propagation, leading to better video quality in many specific cases. One of the most important features of intra coding is intra prediction, an efficient method to compress intra-coded blocks.

The purpose of intra prediction coding is to find a predicted block, similar to the current block, to compress intra-coded blocks. With intra prediction, the bit usage of the coding residual becomes small. In the history of video coding, MPEG-1, MPEG-2, MPEG-4, H261, H.263, H.264/AVC, and HEVC, are coding standards that implement the intra prediction method to reduce the residual data of videos. Among them, H.264/AVC and HEVC improve intra prediction in two ways, which are not provided in any previous video coding standards such as MPEG-1, MPEG-2, MPEG-4, H261, and H.263. One of most important improvements is that in these two video coding standards, intra prediction works in the spatial domain, while, in most previous standards, intra prediction work only in the transform domain. More specifically, H.264/AVC and HEVC are capable of predicting the linear block of the current block using the bounding pixels in the spatial domain. However, in MPEG-1, MPEG-2, and H261, only the DC (Direct Current) coefficient is predicted, and, in MPEG-4 and H.263, only the DC and partial AC (Alternative Current) coefficients are predicted. Another improvement of H.264/AVC and HEVC lies in the block size. The block size of intra prediction in previous video coding standards is a fixed value, e.g. 8×8, but H.264/AVC and HEVC have variable block sizes. H.264/AVC has variable block sizes, e.g. 16×16, 8×8, and 4×4, and the block sizes of HEVC's intra prediction, called Prediction Unit (PU), are 64×64, 32×32, 16×16, 8×8, and 4×4.

Since many intra prediction modes have been adopted by H.264/AVC, the coding performance of H.264/AVC is superior to previous

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

* Corresponding author at: Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan.

http://dx.doi.org/10.1016/j.jvcir.2015.01.010
1047-3203/

video coding standards. However, in intra coding, more bits are needed to encode the blocks when compared to inter coding. In one study [2], the bit usage of intra prediction mode information is 12% of the total intra coding bits. In order to reduce the bit usage result from the encoding of intra modes, Jia et al. utilized the neighboring blocks' intra prediction mode and the border pixel smoothness to select the most probable mode (MPM) in intra prediction [2]. Zhang et al. proposed a context-adaptive coding scheme based on the Markov random field for encoding intra prediction mode information [6]. These approaches reduce the bit usage resulted from encoding the mode information of intra prediction. Another way of reducing intra coding bit usage is to develop a more accurate intra prediction mode. Zhang et al. proposed three additional prediction modes to increase prediction precision [7]. Wang et al. proposed a weighted cross prediction method that replaces the DC mode [8]. Yu et al. proposed an intra prediction mode based on the motion estimation technique [9]. However, all of these methods still require many bits to transmit the coding table or the motion vector (MV). A template matching method is another technique applied to intra prediction. In [10], an "L"-shape template is used to find the best 2×2 prediction block. Tan et al. improves this method in [11] by applying a directional template and averaging predictors. Lan et al. uses the reconstructed region, both in the intra prediction and in the adaptive transform, to reduce the residual bits of a video [12]. Gu et al. applies different templates according to the intra prediction mode to further reduce the residual in the intra lossless coding [13]. Although some template matching methods do not require any additional bits [10,12], these methods have the problem of requiring a huge amount of computation because of the large number of pixels in a template. Hence, an adaptive template matching technique is applied in the proposed method by restricting the number of pixels in a template. As the method with the most representative pixels in a template, the number of pixels in a template is reduced. Thus, a huge computation is saved. Our proposed method applies a finite state machine to both the encoder and decoder for bitrate reduction. In our method, if the prediction block is constructed based on the prediction of the finite state machine, it is not required to transmit extra bits to indicate the position of the most similar block. Thus, the bitrate is reduced significantly. In this paper, we present a substantial theoretical analysis and experiment of the proposed intra prediction modes.

The rest of this paper is organized as follows: Section 2 introduces the intra prediction mechanisms of H.264/AVC and HEVC. Section 3 reviews the backgrounds of an intra prediction method based on the motion estimation technique, and the finite state machine in vector quantization (VQ) coding. Section 4 describes the proposed method, which is based on a finite state machine. Experimental results are presented in Section 5 to demonstrate the effectiveness of the proposed method, while also presenting a comparison between our method and existing methods. Concluding remarks and a discussion of future work are given in Section 6.

2. The intra prediction of video coding standards

2.1. 264/AVC intra prediction

Intra prediction assumes that correlations exist between neighboring blocks. Pixels from the upper and the left of the neighboring encoded blocks are first extrapolated to construct a H.264/AVC prediction block. An intra-coded macroblock (MB) in H.264/AVC with a size of 16×16 is divided into three types: one 16×16 coded block, four 8×8 coded blocks, and sixteen 4×4 coded blocks, which are called Intra16×16, Intra8×8, and Intra4×4, respectively. Four prediction modes are considered in Intra16×16 (Fig. 1(a)), and both Intra8×8 and Intra4×4 have nine prediction modes used to predict the current block (Fig. 1(b)).

To evaluate which mode is the best, rate-distortion optimization (RDO) through a Lagrange multiplier is used for comparisons. Therefore, the best prediction of a MB for the block currently being encoded is selected from Intra16×16, Intra8×8, and Intra4×4 according to their Lagrange costs. The major advantage of intra prediction is that it is simple and efficient when constructing homogeneous blocks in a frame; however, some blocks' contents with finer directions or more complex textures cannot be compressed efficiently in the line-based prediction in the intra coding. Moreover, intra coding needs more bits to encode the residual of the current block when compared to inter coding. In order to enhance the coding performance of intra prediction, we use the finite state machine technique to predict the current block from the frame itself.

2.2. HEVC intra prediction

In HEVC, a frame is divided into Largest Coding Units (LCUs) with size 64×64 instead of a MB defined by H.264/AVC. Each LCU is recursively partitioned into four Coding Units (CUs), and forms a quad tree structure. The block size of CU varies between 64×64 and 8×8. Two independent partition structures are defined within a CU, Prediction Unit (PU) and Transform Unit (TU). PUs control the block size of each prediction block, and TUs control the block size of the integer transform. In intra coding, one $2N \times 2N$ CU is divided into one $2N \times 2N$ PU or four $N \times N$ PUs. And each CU contains one or more TU(s) with size(s) ranging from 4×4 to 32×32. Fig. 2 illustrates the relationship between CUs, PUs, and TUs in an intra-coded LCU.

Compared to H.264/AVC, which has nine prediction modes to predict the current block, HEVC extends the prediction modes to 33 directions and two non-directional modes as shown in Fig. 3, increasing prediction precision. The various block sizes and prediction modes results in a huge computation in the selections of both the block sizes and the prediction modes. To reduce the computation, the Hadamard transform absolute difference (HAD) calculates the distortion between the prediction block and the current coded block in each $2N \times 2N$ PU prediction mode. Then, the prediction modes with the minimal HAD is selected as candidates, and the best prediction mode from these candidates is obtained according to the minimal rate-distortion (RD) cost. This process will be also applied to the mode of four $N \times N$ PUs. Finally, the best PU is selected from one $2N \times 2N$ PU or four $N \times N$ PUs by using the RD cost.

3. Background review

3.1. Intra prediction using intra-macroblock motion compensation

Yu et al.'s work provides a high coding performance when compared to previous intra prediction methods. The intra prediction method proposed by Yu et al., compresses a frame efficiently through the motion estimation technique [9]. In video coding, motion estimation improves inter frame coding significantly when predictions are made based on the reference frame, not just the frame itself [14,15]. Motion estimation can successfully remove temporal redundancy and has been adopted in many well-known video coding standards. Based on these previous research results of the temporal domain, Yu et al. achieved a better intra prediction through motion estimation of the spatial domain. For an input block in intra prediction, Yu et al. looks for the most similar block from previously coded data in the current coded frame. Instead of using directional intra prediction as before, Yu et al. encodes a block as a position vector to indicate the relative position of the