J. Vis. Commun. Image R. 22 (2011) 687–696

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Novel intra prediction via position-dependent filtering

Li Zhang [a,*], Xin Zhao [b,c], Siwei Ma [a,d], Qiang Wang [e], Wen Gao [a]

[a] Institute of Digital Media, Peking University, Beijing, China
[b] Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[c] Graduate University of Chinese Academy of Sciences, Beijing, China
[d] State key laboratory of Virtual Reality Technology and Systems, Beijing University of Aeronautics and Astronautics, Beijing, China
[e] Academy of Broadcasting Science, the State Administration of Radio, Film and Television, Beijing, China

A R T I C L E I N F O

Article history:
Received 16 August 2010
Accepted 12 November 2010
Available online 17 November 2010

Keywords:
H.264
Image coding
Intra prediction
Interpolation filtering
Least square method
Mode-dependent directional transform (MDDT)
KTA
Singular value decomposition

A B S T R A C T

This paper presents a novel intra prediction algorithm, named position-dependent filtering (PDF), to improve the intra prediction accuracy. Different from the existing schemes where the samples along one prediction direction are predicted with the same set of filtering coefficients, in the proposed PDF, position-dependent filtering coefficients are employed, i.e., different sets of filtering coefficients are pre-defined for samples with different coordinates in one coding block. For each intra prediction mode, the set of linear filtering coefficients for each position within one block is obtained from off-line training using the least square method. Moreover, to further reduce the algorithm complexity, a simplified PDF (sPDF) is proposed. In sPDF, only a subset of reference samples are used for prediction and the others are discarded because of the minor contribution to intra prediction. The proposed algorithm has been implemented in the latest ITU-T VCEG KTA software. Experimental results demonstrate that, compared with the original KTA with new intra coding tool enabled, up to 0.53 dB of average coding gain is achieved by the proposed method, while applicable computational complexity is retained for practical video codecs.

1. Introduction

Intra prediction (IP) is an important technique in image and video compression to exploit the spatial correlation within one picture. After intra prediction, the differences between estimated and real sample values instead of the original ones are coded and transmitted, known as residues. Good intra prediction techniques can generate residues with low energy, which is of great importance for compression. The development of intra prediction techniques has passed several decades. The basic idea can be tracked back to differential pulse code modulation (DPCM) coding in 1980 [1]. In MPEG-2 [2], transform-domain intra prediction is employed to handle DC coefficients with DPCM. After that, the related research is mainly focused on spatial-domain intra prediction with multiple prediction directions and mode selection [3–5]. One successful example is the intra prediction method defined in H.264/AVC [6]. In H.264/AVC, line-based intra prediction is employed, where the prediction block is created by extrapolating the reconstructed samples surrounding the target block along a specific direction. Moreover, to better capture the local properties of video signal, H.264/AVC employs flexible macroblock partition modes for Intra coding, including 4×4 (Intra4×4), 8×8 (Intra8×8) and 16×16 (Intra16×16). For predicting the luminance component, nine prediction modes (i.e., eight directional modes plus one DC mode) are employed in both Intra4×4 and Intra8×8 modes, and four prediction modes (vertical, horizontal, DC and plane modes) are utilized for Intra16×16. The efficiency of each partition mode is first evaluated by the encoder using a certain criteria, and the one which optimizes the criteria will then be selected for the actual coding.

The recent advances of intra prediction methods for further coding gain in the literature mainly concentrate on two aspects: (1) better utilization of neighboring samples and (2) utilization of global instead of local signal information. For the first aspect, efforts are devoted to investigate more intra prediction modes or better filtering methods. To capture the local information of neighboring reconstructed samples more accurately, 34 prediction modes are employed in angular intra prediction for Intra8×8 [7], and arbitrary directional intra (ADI) for Intra16×16 [8]. The idea of ADI is similar to angular intra prediction except for the different definition of intra prediction direction. In angular intra prediction, the prediction direction is given by the displacement of the bottom row/rightmost column of the block and the reference row/column around the block. While in ADI, the intra prediction direction is determined by a tuple (dx, dy) which defines the line to be used to calculate the reference samples of the current pixel. Constructing several intra prediction subsets by uniform rotation from

* Corresponding author.
E-mail address: li.zhang@pku.edu.cn (L. Zhang).

1047-3203/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2010.11.003

original predictors of H.264/AVC is introduced in [9]. In [10], bidirectional intra prediction is proposed to combine two kinds of the existing intra prediction modes. To get better interpolation filtering coefficients, the available reconstructed region within a template is used to calculate the filtering coefficients [11]. Liu et al. utilized two windows of reconstructed pixels from two previous frames to train the linear prediction coefficients at both encoder and decoder [12]. In [13], Wang proposed a distance-based weighted prediction method to improve the prediction efficiency of DC mode. Later, a weighted cross prediction mode is proposed [14]. With the revised interpolation method, the decorrelation ability is enhanced and higher coding gain can be achieved.

For some usual cases, the image blocks have repeated patterns instead of distinctive direction information. In this case, utilizing the global information instead of the spatial neighboring samples will bring in better coding efficiency. Related research works include intra displacement vector and template matching, etc. [15–20]. In [15], intra displacement compensation (IDC) utilizes an intra displacement vector per block or partition to get the reference samples. Other related works including intra prediction based on template matching using a single template [16], backward-adaptive texture synthesis [17], multiple candidates [18], priority-guided template matching [19] and locally adaptive illumination compensation [20] are proposed and studied.

In our simulations, it is observed that the correlation among different target samples and reference samples shows diversity. Therefore, in this paper, a position-dependent filtering (PDF) approach is proposed to further improve the intra prediction accuracy. In PDF, each target sample is predicted by extrapolating the neighboring reconstructed ones with its own set of linear filtering coefficients. To obtain these sets of filtering coefficients, the least square method is used for off-line training. Furthermore, in consideration of computational complexity, the filtering coefficients are fixed during the encoding or decoding process. Different from the intra prediction method in H.264/AVC, the sets of linear filtering coefficients in our proposed method are designed to be both mode- and position-dependent. That is to say, for different locations of samples within one block, different sets of filtering coefficients are employed. With the proposed method, the intra prediction is further refined and therefore higher prediction accuracy is achieved. Moreover, to further reduce the algorithm complexity, a simplified PDF (sPDF) is proposed. In sPDF, only a subset of reference samples are used for prediction and others are discarded because of their minor contribution to intra prediction.

The remainder of this paper is organized as follows. In Section 2, a brief review of the intra prediction method in H.264/AVC is provided. The proposed PDF are discussed in Section 3, with detailed descriptions including the motivations, PDF algorithm, filtering coefficient matrix derivation process, implementation and simplification. Performance of the proposed method is validated in Section 4 with extensive experimental results and analysis. Finally, in Section 5, the paper is concluded and some expected future work is also presented.

2. Spatial intra prediction in H.264/AVC

In H.264/AVC, for Intra4 $\times$ 4 and Intra8 $\times$ 8 macroblock partitions, nine intra prediction modes are available including one DC mode and eight directional modes. For example, the directional modes in Intra4 $\times$ 4 are illustrated in Fig. 1(a). The prediction block is derived by extrapolating the neighboring reconstructed samples along some direction with fixed coefficients. Fig. 1(b) illustrates 16 samples of one 4 $\times$ 4 block which are predicted by the reconstructed samples of upper and left-hand reference samples of current block. The filtering process to obtain the predictor values is not shown in this figure. In Fig. 1(b), the samples to be predicted in one 4 $\times$ 4 block and their reference samples are labeled as C_i $(0 \leqslant i \leqslant 15)$ and P_j $(0 \leqslant j \leqslant 12)$, respectively. For brevity, let $\vec{P}$ denote a vector containing the reconstructed reference sample values (e.g., P_0 to P_{12} form a 13 $\times$ 1 column vector $\vec{P}$). Let $\hat{\vec{C}}^k$ denote a vector containing the predicted sample values of the current coded block for a specific spatial IP mode k (e.g., predictor values $\hat{\vec{C}}^k$ of all samples C_i form a 16 $\times$ 1 vector $\hat{\vec{C}}_i^k$). The filtering coefficient matrix W^k corresponding to the filtering coefficient sets can be defined to specify the IP mode. Here, W^k may be expressed as follows:

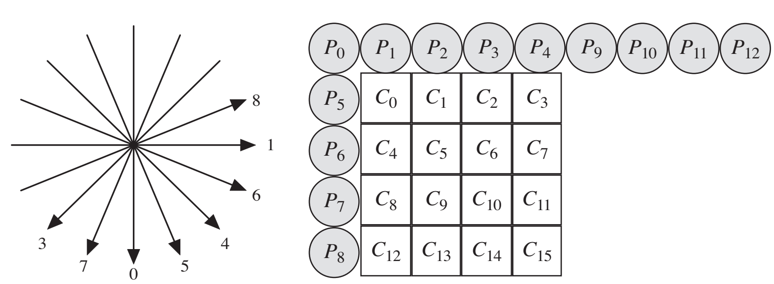
$$W^k = \begin{bmatrix} w^k_{0,0}, & w^k_{0,1}, & \dots, & w^k_{0,11}, & w^k_{0,12} \\ w^k_{1,0}, & w^k_{1,1}, & \dots, & w^k_{1,11}, & w^k_{1,12} \\ \vdots & \vdots & & \vdots & \vdots \\ w^k_{14,0}, & w^k_{14,1}, & \dots, & w^k_{14,11}, & w^k_{14,12} \\ w^k_{15,0}, & w^k_{15,1}, & \dots, & w^k_{15,11}, & w^k_{15,12} \end{bmatrix}. \quad (1)$$

By doing the following multiplication, we can obtain the predicted value vector $\hat{\vec{C}}^k$ as follows:

$$\hat{\vec{C}}^k = W^k \cdot \vec{P}. \quad (2)$$

That is,

$$\begin{bmatrix} \hat{C}^k_0 \\ \hat{C}^k_1 \\ \vdots \\ \hat{C}^k_{14} \\ \hat{C}^k_{15} \end{bmatrix} = \begin{bmatrix} w^k_{0,0}, & w^k_{0,1}, & \dots, & w^k_{0,11}, & w^k_{0,12} \\ w^k_{1,0}, & w^k_{1,1}, & \dots, & w^k_{1,11}, & w^k_{1,12} \\ \vdots & \vdots & & \vdots & \vdots \\ w^k_{14,0}, & w^k_{14,1}, & \dots, & w^k_{14,11}, & w^k_{14,12} \\ w^k_{15,0}, & w^k_{15,1}, & \dots, & w^k_{15,11}, & w^k_{15,12} \end{bmatrix} \cdot \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_{11} \\ P_{12} \end{bmatrix}. \quad (3)$$



(a) Directional intra modes (b) Boundary and inside samples

Fig. 1. Spatial intra4 $\times$ 4 prediction in H.264/AVC.