Improved Intra Prediction Coding Scheme Based on Minimum Distance Prediction for H.264/AVC*

Qingbo WU†(1), Nonmember, Linfeng XU†, Student Member, and Zhengning WANG†, Nonmember

SUMMARY In this letter, we propose a novel intra prediction coding scheme for H.264/AVC. Based on our proposed minimum distance prediction (MDP) scheme, the optimal reference samples for predicting the current pixel can be adaptively updated corresponding to different video contents. The experimental results show that up to 2 dB and 1 dB coding gains can be achieved with the proposed method for QCIF and CIF sequences respectively.

key words: intra prediction, H.264/AVC, video coding

1. Introduction

Along with the series of video coding standards released, video broadcasting and communication technologies have made great progress in recent decades. Compared with previous video compression coding standards, the H.264/AVC [1] directional intra prediction (DIP) coding can reduce the spatial redundancy of adjacent blocks efficiently within a intra frame. In H.264/AVC, 9 prediction modes are used for intra $4 \times 4$ and intra $8 \times 8$ blocks, including the DC mode and the 8 directional modes as shown in Fig. 1 (a). The location relationship for intra prediction is shown in Fig. 1 (b), where the gray squares represent the reconstructed reference samples and the white squares represent the pixels to be predicted.

The H.264/AVC standard has attracted many researchers’ interesting in the past years. Some works focus on speeding up the intra coding module, like [2]–[5]. And some works introduce the perceptual coding or region of interest coding with the computer vision methods like [6]–[8]. Recently, more and more researchers work on improving the RD performance of the H.264/AVC. In DIP scheme, the fixed weights and reference samples can’t achieve optimal prediction for different sequence contents. A number of methods have been proposed to improve the DIP scheme. Yan et al. added 9 novel bi-directional intra prediction (BIP) modes [9]. Kim et al. improved the prediction accuracy of most probable mode to reduce the overhead for signaling the intra prediction modes [10]. The context-adaptive pixel based prediction (CAPBP) algorithm was proposed [11] which adaptively updated the weights by means of solving a least-squares equation. The position-dependent linear intra prediction (PDLIP) approach was proposed [12] to modify the weights of several existing intra prediction modes according to the sequence content. In most existing literatures, the improvement is based on designing better weights for fixed reference samples. There has been few discussions about how to choose optimal reference pixels.

In this letter, an improved intra prediction scheme based on minimum distance prediction is proposed to adaptively select optimal reference samples according to sequence contents. The proposed scheme provides a novel intra prediction mode which not focuses on solving optimal weights, but tries to select the optimal reference samples with the help of a minimum distance matrix (MDM).

The remainder of this letter is organized as follows. Section 2 analyzes the weighted average scheme for directional intra prediction, and Sect. 3 describes the implementation of the proposed MDP scheme in details. Experimental results are reported in Sect. 4, and the conclusion is drawn in Sect. 5.

2. Analysis of Weighted Sum Scheme for Intra Prediction

2.1 Linear Weighted Average Process for DIP

In traditional DIP process, each pixel within current block under some intra prediction mode $k$ will be predicted as the linear weighted average of the surrounding reconstructed samples as shown in Fig. 1 (b). Without loss of generality, we only discuss the case in intra $4 \times 4$ blocks, and it can be easily extended to other block sizes. For the $i$th pixel under prediction with intra mode $k$ in current block, the estimation can be represented as

$$\hat{I}_k(i) = \sum_{m=0}^{M-1} w_j(m) \times P_j(m)$$

(1)
where \( \hat{I}_k(i) \) is the estimated intensity value with \( 0 \leq i \leq 15 \) as shown in Fig. 1(b), and \( k \) represents the index of intra mode selected. \( M \) represents the number of selected reference samples for current pixel under prediction. \( P_i(m) \) represents the intensity value of the \( m \)th neighboring reconstructed sample, and \( w_i(m) \) represents its weight.

In the H.264/AVC standard, the weights designed for each intra prediction mode in \( 4 \times 4 \) blocks are all normalized as

\[
\sum_{m=0}^{M-1} w_i(m) = 1, \quad w_i(m) \geq 0
\]

(2)

where the larger weight would be assigned to the reference sample which is close to the predominant direction specified by intra prediction mode.

Unlike the DIP scheme which selects reference samples via directional characteristic, we determine the optimal candidate samples in terms of intensity distance which can be defined as

\[
E_i(m) = |I(i) - P_i(m)|
\]

(3)

where \( I(i) \) represents the actual intensity value of the pixel under prediction.

Apparently, the optimal reference samples would have minimum distance relative to current pixel to be predicted. And the index \( MD_i \) of the optimal reference sample for the \( i \)th pixel to be predicted can be found by solving the following equation

\[
MD_i = \min_{m \in S_{Idx}} E_i(m)
\]

(4)

where \( S_{Idx} \) represents the index set of all available reference samples.

In DIP scheme, the number of selected reference samples is limited to \( M \in \{1 \sim 3\} \) for directional intra modes. If only one sample is selected, the reference sample with minimum distance will produce the optimal prediction. If there are more than one samples selected, we can summarize their distributions as following two cases: (a) Reference samples are on both sides of current pixel intensity value; (b) Reference samples are on the same side of current pixel intensity value.

For clarity, the relationship between two reference samples and the predicted value has been shown in Fig. 2. From the weighted average scheme in (1), we know that the predictor \( \hat{I}_k(i) \) is between reference samples’ max value and min value. And the predicted value will be closer to the sample with greater weight. So, in case (a), if the greater weight is assigned to the sample which is far different from current pixel, the sample with minimum distance will be closer to current pixel’s actual intensity value comparing with the predicted value as shown in Fig. 2(a). If the weight is assigned to appropriate reference sample, the weighted average predicted value may be better. In addition, in case (b), we can find that the sample with minimum distance is always closer to current pixel comparing with the predicted value. As shown in Fig. 2(b), we can deduce that \( \min(a0, u1) \leq v \leq \max(a0, u1) \) from the weighted average scheme. According to the definition in (4), we know that \( u2 \leq \min(a0, u1) \). Finally, we can conclude that \( u2 \leq v \). Apparently, the prediction from single sample with minimum distance shows the potential in enhancing the prediction accuracy of DIP scheme.

2.2 The MDM Construction Process

To efficiently implement the minimum distance based prediction method, we introduce an indexing structure named MDM which indicates the reference location of a neighboring sample with minimum distance relative to current pixel. As shown in Fig. 3, the following three-step scheme is used to construct the MDM for each intra block. Firstly, we compute all distances between each pixel in intra \( 4 \times 4 \) block and the available reference samples to form a \( 16 \times 13 \) distance matrix where the row corresponding to pixels under prediction and the column corresponding to the reference samples. Secondly, we need to find the indexes of reference samples with minimum distance for each row of the distance matrix. Thirdly, we will return the reference indexes of each row to a \( 4 \times 4 \) minimum distance matrix corresponding to the rows’ original position in intra \( 4 \times 4 \) block.

After obtaining the MDM, we can modify the prediction value in (1) as

\[
\hat{I}_k(i) = P_i(MD_i)
\]

(5)

![Fig. 2](image)

Fig. 2 Relationship between reference samples and the predictor: (a) shows the location relationship in case (a), (b) shows the location relationship in case (b). The symbols \( u0 \sim u2 \) represent the distance between the actual value and the reference samples, and symbol \( v \) represents the difference between the actual value and the predicted value. The scale on the left indicates the intensity level range for 8-bit case.

![Fig. 3](image)

Fig. 3 MDM construction process: the red line connects reference samples involved in this process, the elements in \( 16 \times 13 \) distance matrix like \( a \) represents the distance between a and A.
where $MD_j$ represents the index of the reference sample labeled in MDM.

To evaluate the efficiency of the modified prediction scheme, we use different DIP modes and MDM scheme to implement intra prediction for an intra $4 \times 4$ block, and the prediction distortions measured by the sum of square errors (SSE) are compared in Fig. 4. It can be seen that the SSE produced by MDM scheme is far less than the other intra prediction modes.

Even though a better prediction can be achieved with the help of MDM, we must expend lots of bits for the MDM. In intra $4 \times 4$ block, at least 4 bits are needed to indicate an index in MDM. Too much overhead which will counteract the coding gain from better prediction. So we propose the MDP scheme to predict MDM.

3. Implementation of MDP in H.264/AVC Framework

The framework of our proposed MDP scheme has been diagramed in Fig. 5. In this process, the first frame will still use DIP scheme with fixed weights and reference samples, and the initial MDM will also be constructed here. For the subsequent frame, the MDP scheme will be added as an extra mode along with existing DIP modes. The rate-distortion optimization method is employed to determine the best intra prediction mode with minimum RD cost.

The MDP scheme will be iteratively implemented between the previous reconstructed frame and current frame. Although the adjacent frame is utilized to update the MDM for current frame like [12], the reference samples are still located in current frame and there are no quantization distortion propagation between the consecutive frames.

For clarity, the intra prediction process with MDP scheme is shown in Fig. 6. We define the MDM of the $n$th $4 \times 4$ block in the $l$th frame as $M_l(n)$ and the MDM in previous frame as $M_{l-1}(n)$. On the assumption that the MDM in current frame is highly correlated with the one in previous frame, we directly predict $M_l(n)$ from $M_{l-1}(n)$. In addition, $N$ represents the number of $4 \times 4$ blocks in a frame and $L$ represents the number of frames for coding.

In the MDP scheme, we need at least one bit to indicate whether the MDP scheme is used or not in current block. To further reduce the overhead, we reuse the most probable mode. As the most probable mode or our MDP method is selected, the flag of most probable mode flag will be set to ‘1’. And an extra bit of MDP_flag will be sent to indicate whether the MDP scheme is used. If MDP_flag is ‘1’, the extra mode is selected, otherwise the most probable mode is used. If the flag of most probable mode is ‘0’, the MDP_flag will not be needed.

4. Experimental Results

In order to evaluate the performance of the proposed intra coding scheme, we implement our proposed method in H.264/AVC reference software JM10.2 [13]. The DIP and BIP [9] schemes are compared with the proposed MDP method. In our experiment, only luma component in $4 \times 4$ block are taken into account. The first 150 frames are encoded for each sequence. Both the intra only (INTRA) and IPPP only (INTER) simulations are implemented in our experiment. The QP values for I-frame are 22, 27, 32 and 37. For P-frame, the QP values are incremented by 1. The BDP-SNR [14] is employed to evaluate the RD performance. The run-time percentage difference $\Delta T$ (%) is used to evaluate the computational complexity, which can be formulated as

$$\Delta T = \frac{T_{MDP} - T_{DIP}}{T_{DIP}} \times 100(\%)$$

where $T_{MDP}$ and $T_{DIP}$ denote the run-time of the MDP and DIP schemes, respectively.

The results shown in Fig. 7 present the coding performances for some selected sequences on INTRA and INTER simulations. The detailed coding gains of BIP and MDP...
Fig. 7  RD curves for INTRA and INTER simulations.

Table 1  Comparison of the performance between BIP and MDP schemes in term of BDPSNR (dB) and BDBR (%).

| Sequence       | BIP NPERS  | BIP PSNR  | BIP PSNR%  | MDP NPERS  | MDP PSNR  | MDP PSNR%  | BDBR (%) |
|----------------|------------|-----------|------------|------------|-----------|------------|----------|
| QCIF           | 1.05       | 0.43      | 0.30       | 0.43       | 0.30      | 0.30       | -0.72    |
| dog           | 1.05       | 0.43      | 0.30       | 0.43       | 0.30      | 0.30       | -0.72    |
| car           | 1.05       | 0.43      | 0.30       | 0.43       | 0.30      | 0.30       | -0.72    |
| Avg           | 1.05       | 0.43      | 0.30       | 0.43       | 0.30      | 0.30       | -0.72    |

table data are shown in Table 1. It can be seen that the proposed MDP scheme can achieve significant coding gains relative to the DIP scheme. Especially for the sequence pairs, up to 2.06 dB and 0.45 dB coding gains have been achieved on the INTRA and INTER simulations, respectively. In addition, the complexity investigation for our proposed MDP scheme is shown in Table 2. For both of the INTRA and INTER simulations, only a slight complexity increase is induced in the encoder. In the decoder, the complexity increase is higher. Since there are no mode decision or motion estimation processes in the decoder, the complexity of the decoder is more sensitive to our MDP scheme.

5. Conclusion

In this letter, an improved intra prediction scheme based on minimum distance prediction was proposed for high-performance video coding. The proposed MDP scheme improves the intra prediction accuracy by selecting the optimal reference samples according to different sequence contents. Experimental results demonstrate that the proposed scheme is significantly superior to the traditional DIP scheme in H.264/AVC.

References

[1] ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4) AVC, Advanced Video Coding for Generic Audiovisual Services, 2005.
[2] D. Quan and Y.S. Ho, “Fast intra prediction mode decision for h.264/avc,” IEICE Trans. Inf. & Syst., vol.E93-D, no.7, pp.2012–2015, July 2010.
[3] J.H. Kim, B.G. Kim, and C.S. Cho, “Fast intra-mode prediction algorithm in h.264/avc video coding,” IEICE Trans. Inf. & Syst., vol.E90-D, no.8, pp.1320–1323, Aug. 2007.
[4] H. Li, K.N. Ngan, and Z. Wei, “Fast and efficient method for block edge classification and its application in h.264/avc video coding,” IEEE Trans. Circuits Syst. Video Technol., vol.18, no.6, pp.756–768, June 2008.
[5] Z. Wei, K.N. Ngan, and H. Li, “An efficient intra mode selection algorithm for h.264 based on edge classification and rate-distortion estimation,” Signal Process., Image Commun., vol.23, no.9, pp.699–710, Oct. 2008.
[6] H. Li, G. Liu, Z. Zhang, and Y. Li, “Adaptive scene-detection algorithm for vbr video stream,” IEEE Trans. Multimed., vol.6, no.4, pp.624–633, Aug. 2004.
[7] H. Li and K.N. Ngan, “Saliency model based face segmentation in head-and-shoulder video sequences,” J. Visual Communication and Image Representation, vol.19, no.5, pp.320–333, 2008.
[8] W. Luo, H. Li, G. Liu, and K.N. Ngan, “Global salient information maximization for saliency detection,” Signal Process., Image Commun., vol.27, no.3, pp.238–248, March 2012.
[9] Y. Ye and M. Karczewicz, “Improved h.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning,” Proc. IEEE ICIP, San Diego, CA, USA, pp.2116–2119, 2008.
[10] D.Y. KIM, D.K. KIM, and Y.L. LEE, “A new method for estimating intra prediction mode in h.264/avc,” IEICE Trans. Fundamentals, vol.E91-A, no.6, pp.1529–1532, June 2008.
[11] Y. Zhang, L. Zhang, S. Ma, D. Zhao, and W. Gao, “Context-adaptive pixel based prediction for intra frame encoding,” Proc. IEEE ICASSP, Dallas, TX, USA, pp.898–901, 2010.
[12] L. Zhang, S. Ma, and W. Gao, “Position dependent linear intra prediction for image coding,” Proc. IEEE ICIP, pp.2877–2880, Hong Kong, 2010.
[13] “H.264/avc reference software jnm 10.2.” http://iphome.hhi.de/suehring/tml/download/old_jnm/, 2011.
[14] G. Bjontegaard, “Calculation of average psnr differences between rd-curves,” 13th VCEG-M33 Meeting, Austin, Texas, USA, 2001.