Mode-Dependent Pixel-Based Weighted Intra Prediction for HEVC Scalable Extension

Tang Kha Duy Nguyen*[a], Chun-Chi Chen[a]
[a]Department of Computer Science, National Chiao Tung University, Taiwan

ABSTRACT

The current draft scalable extension to HEVC offers two approaches, RefIdx and TextureRL, for performing inter-layer prediction. In the framework of TextureRL, this paper presents a mode-dependent pixel-based weighted intra prediction scheme for coding the enhancement layer (EL). The scheme first decomposes the EL intra prediction and the collocated base layer reconstructed block into their respective DC and AC components and then computes a weighted sum of both to form a better prediction signal using a pixel-based weighting scheme. The weighting factors to associate with different components are obtained by a least-squares fit to the training data. It was observed that they depend strongly on the EL's intra prediction mode and prediction block size, but are less dependent on QP settings. The experimental results show an average BD-rate savings of 1.0% for the AI-2x configuration and 0.5% for AI-1.5x over the SHM-1.0 anchor.

Keywords: SHVC, Scalable Video Coding, Intra Prediction, TextureRL.

1. INTRODUCTION

Scalability extension to HEVC (called SHVC) [1][2] is also realized by the layered approach as its predecessor, H.264/SVC [3]. The base layer (BL) contains the most fundamental information to ensure a minimum guaranteed decoded quality, while the enhancement layers (ELs) produce improved quality (in terms of frame rate, resolution and reconstruction SNR) when combined with the BL for decoding. Since the input video for all layers are from the same content, there is a strong correlation between the EL and the BL. Therefore, the number of bits required to code the EL can be significantly reduced by exploiting that correlation between the layers. As an example, in H.264/SVC, the redundancies between layers are reduced by employing various inter-layer prediction (ILP) mechanisms, such as inter-layer texture, motion, and residual prediction.

In the present test model of SHVC, the ILP mechanisms can be realized by two different approaches, the RefIdx and the TextureRL [4]. In the RefIdx approach, the BL reconstructed picture is included in the reference picture lists, so that the inter-layer motion or texture prediction can be achieved by simply using the existing inter-frame prediction mechanisms. This approach has the benefit of being able to reuse most of the single layer HEVC design, except for a few high-level (e.g. slice level) syntax changes, for encoding and decoding the EL. On the contrary, the TextureRL approach allows changes at the low-level (e.g. CU- and/or PU-level) syntax and decoding process, in hopes of getting a higher coding gain. Without being constrained to use inter-frame prediction mechanisms, this approach has more flexibility in forming ILP. For example, a new prediction mode may be created at the PU level to better utilize both the BL and EL information for prediction. The price paid for this flexibility is however an EL codec that is less compatible with HEVC. It was shown that the RefIdx approach can already achieve most of the gain provided by the TextureRL [5][6], making the latter a less favorable option. But, many believe the flexibility of the TextureRL has not been employed to the best advantage.

Aiming to explore the potential of the TextureRL approach in forming a better ILP, Lainema et al. [7] devises the notion of Intra DC Correction (IDCC) to improve the EL intra predictor by replacing its mean value (referred hereafter to as DC value) with that of the collocated block in the BL, which is motivated by the observation that the collocated BL block often gives a better estimate of the DC value of the input block than the EL intra predictor. However, the IDCC algorithm fails to estimate the value for those pixels that are located adjacent to the reference pixels due to the fact that those pixels are well predicted by the conventional intra prediction direction. C. K. Kim [8] introduces a Weighted Intra Prediction (WIP) scheme to linearly combine the EL intra predictor with the BL reconstructed block based on a pixel-based weighting scheme. The major issue of the WIP scheme is that it employs a single weighted scheme for all intra prediction directions.

Sixth International Conference on Digital Image Processing (ICDIP 2014), edited by Charles M. Falco,
Chin-Chen Chang, Xudong Jiang, Proc. of SPIE Vol. 9159, 915925 · © 2014 SPIE
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2064639

We attempt to combine the merits of these two proposed notions to form a prediction scheme that can improve further on prediction efficiency, and thus a mode-dependent pixel-based weighted intra prediction scheme (MPWIP) for coding the EL is proposed. Preliminary results show that the MPWIP provide 0.5-1.0% BD-rate reduction with respect to SHM 1.0.

The rest of this paper is organized as follows: Section 2 presents our proposed algorithm, including its concept of operations and the mathematical model. Section 3 presents the weighting schemes for different components, and how their weighting should be adapted in response to the change in coding parameters. Section 4 provides experimental results of the proposed algorithm. Finally, Section 5 concludes this work with a summary of our findings.

2. PROPOSED ALGORITHM

This section presents the proposed MPWIP starting by the introduction to its concept of operations and then the employed mathematical model. As we shall see, the weight value associated with each component is designed to be a function of the prediction pixel's position within the block, which we refer hereafter to as the weighting function.

2.1 Concepts of operation

The ideas of the proposed scheme are essentially a combination of the two prior works [7] and [8]. As depicted in Figure 1, the texture information of the EL intra prediction and BL reconstructed block are first decomposed into the DC and AC components, where the DC components are computed as a prediction block with all its pixels taking the average value of the input block, i.e., the EL intra predictor or the BL reconstructed block, and the AC components are formed from the residual signals produced by subtracting the DC components from their respective inputs. Each of these components is then weighted by a separate pixel-based weighting scheme and the results are added together to form the final predictor.

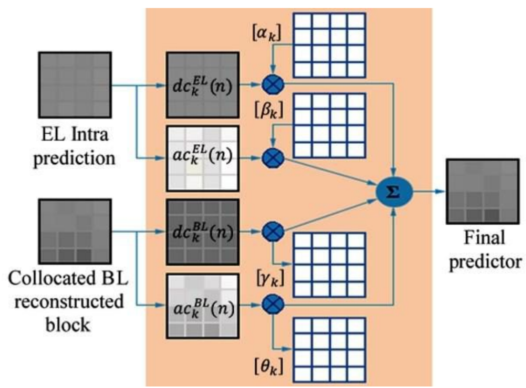

Figure 1. Proposed Pixel-based Weighted Intra Prediction scheme.

2.2 Iterative Least-Squares Solution

Obviously, how different components are weighted has a crucial effect on the resulting prediction performance. This research proceeds in the hope that we can find a set of weighting functions where the resulting prediction residual can be minimized. This problem can be solved using the well-known Least-Squares (LS) method.

Specifically, $\mathbf{a}_k = [a_k(1) a_k(2) \ldots a_k(n)]^T$ and $\mathbf{b}_k = [b_k(1) b_k(2) \ldots b_k(n)]^T$ represent the predictor values at pixel k that are extracted from the EL intra predictors and the BL reconstructed blocks, respectively, in n collected data from the training process. Similarly $\mathbf{ac}_k = [ac_k(1) ac_k(2) \ldots ac_k(n)]^T$ and $\mathbf{dc}_k = [dc_k(1) dc_k(2) \ldots dc_k(n)]^T$ denote respectively the corresponding values from the AC and DC components. Thus, we have

$$\mathbf{a}_k = \mathbf{ac}_k^{EL} + \mathbf{dc}_k^{EL} \quad (1)$$

$$\mathbf{b}_k = \mathbf{ac}_k^{BL} + \mathbf{dc}_k^{BL} \quad (2)$$

With these, the final prediction signal $\mathbf{p}_k = [p_k(1) p_k(2) \ldots p_k(n)]^T$ at pixel k can be written as:

$$\mathbf{p}_k = \mathbf{X}_k \mathbf{w}_k \quad (3)$$

Where $\mathbf{X}_k = [\mathbf{ac}_k^{EL} \mathbf{dc}_k^{EL} \mathbf{ac}_k^{BL} \mathbf{dc}_k^{BL}]$ is a set of basis functions and $\mathbf{w}_k = [\alpha_k \beta_k \gamma_k \theta_k]^T$ is a weight vector whose elements are the weight values to associate with the four components for prediction at pixel k. Specifically, the weight vector represents the weighting scheme from the perspective of a single pixel, describing how the corresponding samples from different components contribute to estimating a current pixel's intensity.

With reference to the notations above, we further denote $\mathbf{o}_k = [o_k(1) o_k(2) \dots o_k(n)]^T$ as the target pixels at position k in n collected blocks, whose intensity values are to be estimated. The problem of determining the optimal weight vector $\mathbf{w}_k^*$ in the least-squares-error sense can then be formulated as follows:

$$\mathbf{w}_k^* = arg \min_{\mathbf{w}_k} (\mathbf{o}_k - \mathbf{X}_k \mathbf{w}_k)^2 \tag{4}$$

From the Linear Algebra theorem, it has the closed-form solution

$$\mathbf{w}_k^* = (\mathbf{X}_k^T \mathbf{X}_k)^{-1} \mathbf{X}_k^T \mathbf{o}_k \tag{5}$$

By varying the index k and repeating the same process, we can obtain the weight vectors for different pixel positions. Thus the weighting function, a function of the prediction pixel's position within the block representing the weighting scheme from the perspective of a single component (e.g. the AC or the DC component at the BL or at the EL), can be constructed for all components.

In order to collect the data for the basis functions in (3), the training process is introduced. First, the proposed algorithm will be applied to produce a new prediction mode and this mode has to compete with all conventional modes in the rate-distortion optimization process at the EL to find the best mode with lowest rate distortion cost. Then, those blocks coded in the proposed algorithm will be used to compute the optimal weighting functions with respect to (5). The records show that the optimal values can be achieved at the fifth iteration of training process. It is noted that the sequences of the training set differ from those of the test set; therefore, the obtained weighting functions used to find the bit rate savings on the test sequences would be referred to as the weighting functions that are resulted from the training process.

3. ANALYSIS OF THE WEIGHTING FUNCTIONS

To gain a better understanding of how different components should be weighted in forming a better predictor, this section provides an in-depth analysis of the weighting functions with different components against 1) prediction mode taken by the EL, 2) QP setting of BL and EL, and 3) prediction block size.

3.1 Effect of intra prediction mode

This subsection investigates the effect of the intra prediction mode on the weighting function. Currently, our weighted scheme is restricted to the cases where the EL predictor is produced with Horizontal, Vertical, DC or Planar mode. Figure 2 shows the weighting functions for Horizontal mode for a 16x16 block. It can be seen that those associated with the components from the same layer have a similar waveform, although their magnitudes differ considerably. Moreover, the weight value for the DC component of the EL is seen to be mostly lower than that of the BL, which justifies the IDCC algorithm's substitution of the BL's DC value for the EL's. We can further observe that the weighting functions are mode dependent due to the fact that the waveforms vary according to the direction of the intra prediction mode.

3.2 Effect of QP setting

This subsection studies the effect of the QP setting (QP values assigned to the BL and the EL) on the weighting functions. Plotted in Figure 3 is the weighting function of the EL's DC component along the slice of Y=10 for various QP settings (with the QP_{BL} ranging from 22 to 34, increasing by 4, and delta QP (DQP) = 0 or 2, DQP is the difference in QP values between the QP_{EL} and the QP_{BL}); the results shown correspond to the Horizontal mode.

As expected, all these weighting functions decrease with the increasing X value because of the nature of Horizontal prediction and this is a result that we have seen before. Of more interest is the observation from the weight values of the QP settings in the same set that even though the difference between the smallest and the largest QP settings are considerable (with the value of 22 for the first and 34 for the latter case), the variance in the magnitude of those waveforms is insignificant. From extensive experiments with all possibilities of QP setting combinations (details are discussed in section IV), we found that the effect of the QP setting on the weighting functions in terms of bit rate savings is insignificant. Therefore, these observations can lead to an opportunity to unify the weighting functions for the QP settings of all test sequences in the common test conditions.

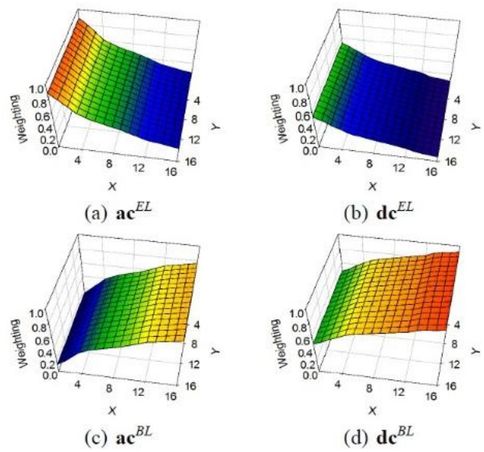

Figure 2. Waveforms of weighting functions for AC and DC components of the EL and the BL obtained from Horizontal mode, block size of 16x16.

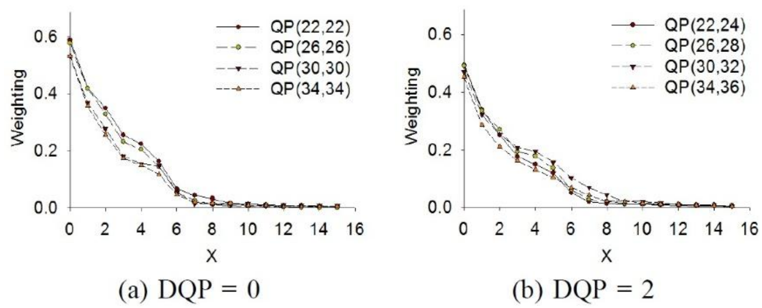

Figure 3. The curves of DC component in the EL of Horizontal mode on different QP settings in two sets of coding configurations. (a) DQP = 0: $QP_{EL} = QP_{BL}$, (b) DQP = 2: $QP_{EL} = QP_{BL} + 2$.

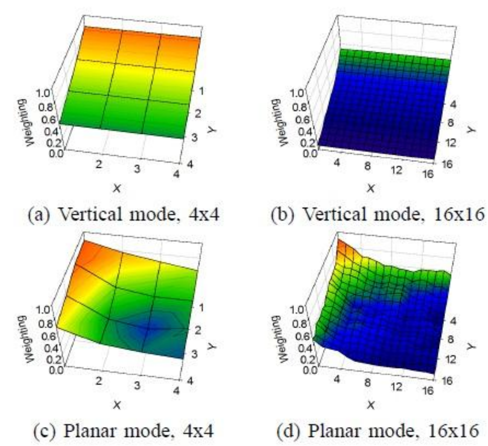

Figure 4. Waveforms of weighting functions for DC component at the EL of Vertical and Planar mode of different block size levels.

3.3 Effect of prediction block size

This last investigation shows how the prediction block size affects the weighting functions. The effect of the prediction block size is somewhat expectable (the higher weight value for the EL's components will be given to the smaller block size); however, the texture information in the BL perhaps varies the expectation. As a result, the analysis on the effect of the prediction block size is necessary. To this end, Figure 4 contrasts the weighting function of the EL's DC component for 4x4 and 16x16 block sizes. Results are given for Vertical and Planar modes. As expected, the simulation results show that the EL predictor tends to have a higher weight value across the entire prediction block when the block size is smaller. This is understandable given that directional intra prediction usually performs more efficiently with a smaller block size and that the EL reference pixels are subject to less coding error.

4. EXPERIMENTS

4.1 Implementation and experimental conditions

We have implemented our proposed scheme in SHM1.0 [4] and conducted extensive experiments, following mainly the All Intra (AI) common test conditions [9] to investigate the performance and compared it with the other prior works. In particular, only the results for mandatory tests, in which the base layer is HEVC coded, are presented. Moreover, for the 2-layer spatial scalability, the ratio of the EL's resolution to that of BL is limited to 2x and 1.5x, although in principle any resolution ratios between the layers, including the ratio of 1, can be considered.

The experiments were performed on the test sequences of 2 classes (sequences of class A and B at 2560x1600 and 1280x720 for the EL, respectively). The resolution of the EL is to be fixed while the resolution of the BL varies according to the test conditions (AI-2x and AI-1.5x cases). Moreover, for the test sequences of class A, only the AI-2x test condition is applied. Four QP values were used per sequence, with the EL QP values set to be equal and to an offset of 2 above the BL QP values.

4.2 Coding performance of MPWIP

From Table 1, the overall BD-rate Y gain over the anchor is 1.0% for the AI-2x case and 0.5% for the AI-1.5x case. The coding gain achieved, however, is highly variable over the test sequences. As an example, the smallest gain is in the test sequence ParkScene, with only a 0.3% and 0.1% gain for AI-2x and AI-1.5x, respectively, whereas a much higher improvement is observed in the BasketballDrive sequence, reaching up to 2.4% and 1.5%, respectively. This can be explained by the characteristics of each sequence. The ParkScene is a highly-textured sequence while the BasketballDrive contains more homogeneous regions.

Table 1. Coding Performance of the MPWIP.

		AI HEVC 2x			**AI HEVC 1.5x**		
Test Class	Test Sequence	Y	U	V	Y	U	V
A	Traffic	-0.7%	-0.6%	-0.6%	N/A		
	PeopleOnStreet	-0.6%	-0.7%	-0.4%			
B	Kimono	-0.4%	-0.4%	-0.5%	-0.2%	-0.3%	-0.3%
	ParkScene	-0.7%	-0.7%	-0.7%	-0.1%	-0.2%	-0.3%
	Cactus	-0.7%	-0.7%	-0.7%	-0.3%	-0.6%	-0.6%
	BasketballDrive	-2.4%	-2.7%	-2.7%	-1.5%	-1.6%	-1.3%
	BQTerrace	-1.3%	-1.6%	-1.8%	-0.4%	-0.6%	-0.7%
Overall (EL + BL)		-1.0%	-1.1%	-1.0%	-0.5%	-0.7%	-0.6%

For comparison with the IDCC and WIP algorithms, their respective average BD-rate savings in different test classes and configurations relative to the anchor are shown in Table 2. It can be seen that both schemes offer a much smaller BD-rate savings than our MPWIP. Specially, that of the IDCC ranges from 0.0% to 0.5% in all test cases, with an overall savings of no more than 0.2%. The WIP, although performing relatively better in some sequences, resulted in the higher peak of maximum coding gain, shows a similar performance to the IDCC in terms of BD-rate Y.

Table 2. Comparison of the MPWIP and prior works.

	AI HEVC 2x			**AI HEVC 1.5x**		
Test	IDCC	WIP	MPWIP	IDCC	WIP	MPWIP
Class A	-0.1%	-0.2%	-0.7%	N/A		
Class B	-0.2%	-0.3%	-1.1%	-0.1%	-0.1%	-0.5%
Min.	0.0%	0.0%	-0.3%	0.0%	0.0%	-0.1%
Max.	-0.5%	-0.7%	-2.4%	-0.1%	-0.3%	-1.5%
Overall (EL + BL)	-0.2%	-0.3%	-1.0%	-0.1%	-0.1%	-0.5%

5. CONCLUSION

In this paper, we have introduced a sophisticated algorithm to combine the EL intra predictor and the BL reconstructed block targeted to improve the EL intra prediction in the framework of the TextureRL; and this algorithm does bring a coding gain. The proposed scheme achieves 0.5-1.0% BD-rate reductions with respect to SHM 1.0 and an improvement of up to 0.7-0.8%, when compared with the prior works. In addition, the coding parameters (e.g. the intra prediction mode of the EL intra predictor, the QP setting, and the prediction block size) that affect the weighting scheme were thoroughly analyzed. With respect to these analyses, a simplified version of the proposed design (e.g. Unification to MPWIP simplification) is presented in an attempt to reduce the memory requirement introduced by the original design. It is observed that a significant reduction in terms of memory requirement was obtained with a slightly drop of BD-rate savings.

REFERENCES

[1] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[2] J. Chen, J. Boyce, Y. Ye, and M. M. Hannuksela, "SHVC Working Draft 2," JCTVC-M1008, Apr. 2013.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," IEEE Trans. Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[4] J. Chen, J. Boyce, Y. Ye, and M. M. Hannuksela, "SHVC Test Model 1 (SHM 1)," JCTVC-L1007, Jan. 2013.

[5] X. Xiu, Y. He, Y. He, and Y. Ye, "TEA2: Inter-layer reference picture placement," JCTVC-L0051, Jan. 2013.

[6] J. Zan, "TE2: Results of test 3.1.3 on CU based Intra-BL signaling," JCTVC-L0165, Jan. 2013.

[7] J. Lainema and K. Ugur, "TE3: Results of test 4.1.2 on intra DC correction," JCTVC-L0036, Jan. 2013.

[8] C. K. Kim, "Non-CE1: Weighted intra prediction," JCTVC-M0117, Apr. 2013.

[9] X. Li, J. Boyce, and P. Onno, "Common Test Conditions and Software Reference Configurations for the Scalable Test Model," JCTVC-L1009, Jan. 2013.