INTRA PREDICTION MODE CODING FOR SCALABLE HEVC

D. Bugdayci[⋆] *M. O. Bici*[⋆] *K. Ugur*[†] *M. Gabbouj*[⋆]

[⋆] Tampere University of Technology, Tampere, Finland
[†] Nokia Research Center, Tampere, Finland

ABSTRACT

High Efficiency Video Coding (HEVC) standard introduced an increased number of intra prediction directions in order to improve intra prediction performance by efficiently modeling the directional structures found in typical video contents. Efficient coding of intra prediction mode information is realized through a Most Probable Mode (MPM) list approach. In a scalable system, due to high correlation between the layers, utilization of base layer intra prediction mode can improve coding performance. In this paper, we propose a new intra prediction mode coding algorithm for scalable extension of HEVC where only the difference between the intra prediction modes of base and enhancement layers is coded. We provide experimental results and also a comparison of the proposed algorithm with an MPM list based approach where base layer intra prediction mode is added to the list as the most probable mode. Experimental results show BD-rate gains up to 1.1% in 2x spatial scalability and 0.7% in 1.5x scalability for all intra configuration.

***Index Terms*—** HEVC, video coding, scalable, intra prediction mode, differential coding

1. INTRODUCTION

First version of the new video coding standard, HEVC, is planned to be finalized in January 2013. On the other hand, development of scalable extension of the standard has just begun after the joint call for proposals by ITU-T SG 16 Q.6 (Video Coding Experts Group - VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (Moving Picture Experts Group - MPEG) [1]. HEVC is a hybrid video codec utilizing block-based prediction and transform coding similar to previous video coding standard H.264/AVC. The input video is divided into rectangular blocks that are predicted from either previously decoded frames or from the current frame. After prediction, an integer approximation of discrete cosine transform is applied on the prediction error which is followed by quantization and entropy coding of the transform coefficients. HEVC defines a new structure called Coding Tree Unit (CTU) of sizes 16×16, 32×32 and 64×64. A CTU is the root of the coding tree and it can be split into smaller Coding Units (CU). A CU can have either intra or inter prediction, however it can further split into prediction units (PU) to have distinct prediction parameters for smaller regions. CU also splits into transform units (TU) allowing transform block sizes ranging from 4×4 to 32×32 samples. HEVC introduces many new coding tools to achieve high compression performance while dealing with higher resolutions and to allow higher utilization of parallel architectures. For a complete overview of HEVC and detailed explanation of the specific tools the reader may refer to [2].

Scalable extension of H.264/AVC allows inter-layer texture prediction, residual prediction, prediction of macroblock partitioning and motion parameters for spatial scalability [3]. A study by Yang [4] introduces methods to improve intra prediction in scalable extension of H.264/AVC, such as allowing sub-macroblock level inter-layer prediction or reducing the number of candidate intra prediction modes using upsampled base layer. However, none of the previous work reports the direct usage of intra prediction mode information for an improved coding of enhancement layer intra prediction mode. In this study, we propose a new coding method for scalable extension of HEVC which utilizes differential coding of enhancement layer intra prediction mode information with respect to base layer.

In the following section, first an overview of intra prediction mode coding in HEVC is described followed by the explanation of our proposed algorithm and a discussion about the parsing dependency issue caused by our proposal. Section 3 provides the experimental results and final section concludes the paper.

2. INTRA PREDICTION MODE CODING

Intra prediction mode coding in HEVC is described in detail in [5]. HEVC introduces 33 angular prediction modes (Modes 2 to 34, Figure 1) in addition to planar (Mode 0) and DC (Mode 1) modes for intra coding. Due to increased number of modes (35), efficient coding of intra prediction mode is achieved by using a list based approach. For each prediction unit, the most probable 3 modes are determined and a Most Probable Mode (MPM) list is constructed from these modes. The most probable modes are selected among the intra prediction modes of the neighbouring PUs, more specifically the left and the above neighbours. In order to avoid duplicates in the list, left and above neighbour's intra prediction modes are compared. If the two modes are the same and equal to either Planar or DC modes, then the list is constructed as Planar, DC and Angular (Mode 26 - Vertical) in order. If the two modes are the same and equal to an angular mode, then the list is constructed by this mode and two angular modes that are closest to it. If the intra prediction modes of the left and above neighbours are different from each other, they are inserted in the list and the third mode is set as Planar, DC or Angular (Mode 26 - Vertical) in the mentioned order.

After determining the prediction mode of the current PU, the encoder checks whether this prediction mode is available in the MPM list. If so, only the index in the list is signaled. Otherwise, the current prediction mode is signaled after a fixed length binarization and bypass coding. There are 32 modes outside the MPM list which can be represented by 5 bits properly. Coding of luma intra prediction mode consists of three syntax elements, namely *prev_intra_luma_pred_flag*, *mpm_idx* and *rem_intra_luma_pred_mode*. The syntax element *prev_intra_luma_pred_flag* indicates whether the current prediction mode is in the MPM list or not. When this flag is equal to 0, the syntax element *rem_intra_luma_pred_mode* is coded indicating the prediction mode. When the flag is equal to 1, then the syntax element *mpm_idx* is coded from which the decoder can get the current intra prediction mode by constructing the MPM list.

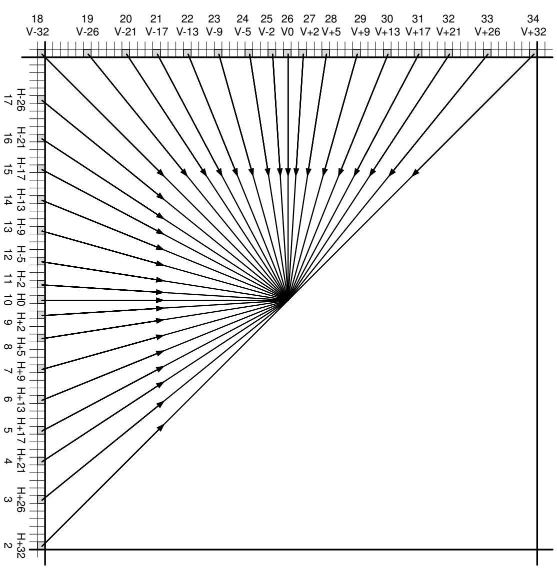

Fig. 1. HEVC angular intra prediction modes ranging from 2 to 34 and the associated displacement parameters.

2.1. Proposed Algorithm

Since base and enhancement layer frames for a given time are the same pictures with different resolution, there is high correlation between their intra prediction modes. For most of the PUs in enhancement layer, intra prediction mode is the same as in the collocated base layer PU. This result is observed by examining the histogram of base and enhancement layer intra prediction mode selections. In order to exploit this correlation, intra prediction mode information of enhancement layer can be coded with respect to base layer. Therefore, we propose to replace the current MPM list based coding of intra prediction mode of enhancement layer with a differential coding scheme. The difference between the intra prediction modes of base and enhancement layer blocks is taken as the value to be coded which ranges from -34 to +34. Histogram of this value for the first frame of BasketballDrive sequence is given in Figure 2. By examining the histogram of various sequences coded with different base and enhancement layer QPs, it is found that the most probable value of the difference is zero followed by ± 1 with similar probability and ± 2 with similar probability. The rest of the difference values are observed to have almost equal (but much lower) probability. In our algorithm, the difference is represented with a sign and a magnitude component and various syntax elements are defined as follows in order to efficiently code this value:

- A binary valued syntax element indicating whether the difference is 0 or not is defined as *luma_intra_pred_equal_base_flag*.
- If *luma_intra_pred_equal_base_flag* is not equal to 1, a binary valued syntax element indicating the sign of the difference is defined as *luma_intra_pred_diff_sign_flag*. Note that this syntax element is only present if *luma_intra_pred_equal_base_flag* takes a value of 0.
- A binary valued syntax element indicating whether the absolute value of the difference is equal to 1 or not is defined as *luma_intra_pred_diff_abs_equal1_flag*.
- If *luma_intra_pred_diff_abs_equal1_flag* is not equal to 1, a binary valued syntax element indicating whether the absolute value of the difference is equal to 2 or not is defined as *luma_intra_pred_diff_abs_equal2_flag*.
- If *luma_intra_pred_diff_abs_equal2_flag* is not equal to 1, a nonbinary valued syntax element indicating the absolute value of the difference minus 3 (which corresponds to the remaining 32 different values ranging from 2 to 34) is defined as *luma_intra_pred_diff_abs_remaining*.

First step of encoding process for CABAC is binarization where nonbinary valued syntax elements are mapped to a binary sequence called bin string [6]. For binary valued syntax elements this binarization step is bypassed. In the above example only the syntax element called *luma_intra_pred_diff_abs_remaining* requires binarization which is done using Fixed-Length (FL) binarization scheme of CABAC. In the

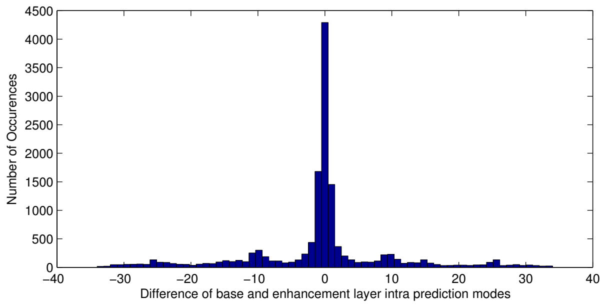

Fig. 2. Histogram of base and enhancement layer intra prediction mode differences for BasketballDrive 2x scalablity

regular coding mode of CABAC, a bin enters the context modeling stage, where a probability model is chosen according to previously coded syntax elements or bins. Then the bin value and the assigned context model are entered to the regular coding engine where the final stage of arithmetic coding takes place together with subsequent model updating. Alternative to regular coding mode, there is also bypass coding mode which is a simplified coding engine that does not use an explicitly assigned model which allows a speedup of the encoding process. In this paper, we have utilized two different coding schemes. In the first coding scheme, syntax elements *luma_intra_pred_equal_base_flag*, *luma_intra_pred_diff_abs_equal1_flag* and *luma_intra_pred_diff_abs_equal2_flag* are coded in regular mode with each of them having their own context model. In the second coding scheme, only the syntax element *luma_intra_pred_equal_base_flag* is coded in regular mode. Restricting the number of bins coded in regular mode is desired to allow speed up and simpler design. In both coding schemes, syntax elements *luma_intra_pred_diff_sign_flag* and *luma_intra_pred_diff_abs_remaining* are coded in bypass coding mode whereas coding order of *luma_intra_pred_diff_sign_flag* may change in order to group the ones coded in bypass mode together to be compatible with the HEVC coding design.

Table 1. Mapping between the intra prediction modes and coefficient scanning order

Intra prediction mode	Scan pattern for 4×4 and 8×8	Scan pattern for 16×16 and 32×32
Angular (6-14)	Horizontal	Diagonal
Angular (22-30)	Vertical	Diagonal
All other modes	Diagonal	Diagonal

2.2. Parsing Dependency

HEVC introduces a new tool called mode dependent coefficient scanning (MDCS) for intra prediction that improves the coding efficiency of 4×4 and 8×8 transform blocks [7]. In MDCS, for 4×4 and 8×8 transform blocks, one of the three possible scanning patterns is selected by using the intra prediction mode information. The mapping between the intra prediction modes and scanning patterns is shown in Table 1. In transform coefficient coding, syntax elements are coded in scan passes which starts from the last significant coefficient and proceeds to the DC coefficient in reverse scanning order. Therefore, the number of coefficients scanned in between varies for each scan pattern. Since for 4×4 and 8×8 blocks, scan index is inferred from intra prediction mode, an enhancement layer that does not have the base layer information may not have the correct intra prediction mode and hence the scan index. This kind of dependency on the base layer is undesired since it makes the enhancement layer undecodable even if it is received correctly due to the unavailability of the base layer information. The dependency is unavoidable with MDCS tool and the current transform coefficient coding design. Therefore, we have disabled the horizontal and vertical scan directions which are defined for small block sizes (4×4 and 8×8) and the scan direction is set to be diagonal for all block sizes.

3. EXPERIMENTAL RESULTS

In this section we provide the performance results of the proposed algorithm for both coding schemes with respect to independent coding of enhancement layer intra prediction mode as in single layer HEVC. The results are compared using percentage of BD-rate [8] where negative numbers indicate gain and positive numbers indicate loss relative to the anchor method. Our anchor software is the HM8.1 based reference software provided by JCT-VC for the development of scalable extension of HEVC. We have implemented our algorithm into this software and also tested the MPM list modification based algorithm present in the software. MPM list based algorithm basically modifies the list generation process so that, base layer intra prediction mode is entered to the list as the most probable mode (into the first location in the list). The rest of the list generation follows the original process by checking the duplicates and filling the list using methods described under Section 2. The results of this algorithm are also provided with respect to the same anchor.

In our experiments we have used a test set consisting of 7 sequences, 2 of which belong to class A and the rest to class B. Class A sequences have a resolution of 2560×1600 and class B sequences have 1920×1080. For 2x scalability both Class A and Class B sequences are used and for 1.5x scalability only class B sequences are used. Base layer sequences are obtained by downsampling the full resolution sequence according to the scalability factor and using the default downsampler defined in HEVC scalable extension call for proposals [1]. Since our porposed algorithm is related to intra prediction, we provide results of all intra (AI) configuration corresponding to GOP size 1. In encoding, we used base layer quantization parameters (QP) 26 and 30 and for each base layer QP, enhancement layer QP is set with a difference of -2, 0, 2 and 4 with respect to base layer QP which results in 8 different encoding scheme for a given sequence and scalablity mode.

Table 2. BD rate (in %) of all intra 2x spatial scalability for coding scheme 1 and 2

	Scheme 1			Scheme 2		
	Y	U	V	Y	U	V
Traffic	-0.5	0.1	0.3	-0.5	0.1	0.4
PeopleOnStreet	-0.8	-0.5	-0.3	-0.8	-0.5	-0.3
Kimono	-0.3	0.1	0.1	-0.3	0.0	0.0
ParkScene	-0.1	-0.1	-0.2	-0.1	-0.1	-0.2
Cactus	-0.9	-0.7	-0.8	-0.8	-0.7	-0.8
BasketballDrive	-1.1	-0.5	-0.5	-1.1	-0.5	-0.5
BQTerrace	-0.9	-0.9	-0.7	-0.9	-0.9	-0.7
Class A average	-0.6	-0.2	0.0	-0.6	-0.2	0.0
Class B average	-0.7	-0.4	-0.4	-0.7	-0.4	-0.4
Average	-0.7	-0.4	-0.3	-0.6	-0.4	-0.3

Table 3. BD rate (in %) of all intra 1.5x spatial scalability for coding scheme 1 and 2

	Scheme 1			Scheme 2		
	Y	U	V	Y	U	V
Kimono	-0.1	0.1	0.2	-0.1	0.2	0.2
ParkScene	-0.1	0.1	0.0	-0.1	0.0	0.0
Cactus	-0.4	0.0	-0.1	-0.4	0.0	-0.1
BasketballDrive	-0.7	0.9	0.7	-0.7	0.9	0.7
BQTerrace	-0.4	0.0	0.1	-0.4	0.0	0.2
Average	-0.3	0.2	0.2	-0.3	0.2	0.2

The results can be compared in two groups. During comparison we consider the enhancement layer gains and losses in Y component primarily, since our algorithm is implemented for luma intra prediction mode only. Tables 2 and 3 provides the percentage BD-rate results of our proposed algorithm with a parsing dependency on base layer where the anchor is independent coding of intra prediction mode of enhancement layer. Scheme 1 correponds to coding scheme 1 described in Section 2.1 where three syntax elements are coded in regular mode whereas scheme 2 corresponds to coding only one syntax element in regular mode. Comparing the average gains brought by these two coding schemes, it can be concluded that from complexity point of view, coding scheme 2 can be chosen as there is marginal loss in coding efficiency in comparison to using coding scheme 1.

Table 4 provides the percentage BD-rate results of our algorithm (coding scheme 2) with MDCS disabled in order to resolve the parsing dependency problem. When we compare the scheme 2 results with and without MDCS, we see the effect of MDCS which is 0.2% loss in Y component in 2x scalability and no loss in 1.5x scalability. Finally, Table 5 shows the results for MPM list based approach with MDCS disabled. Effect of MDCS is similar for this algorithm also, so we did not include results with MDCS enabled here. The results indicate that our algorithm performs slightly better than the MPM list modification. Even though in the worst case, the number of syntax elements to be coded is higher for our algorithm, the probability distribution of the difference value shows that these worst cases happen rarely.

Table 4. BD rate (in %) of all intra for coding scheme 2 with diagonal only scan (no parsing issue)

	2x			1.5x		
	Y	U	V	Y	U	V
Traffic	-0.4	0.4	0.5	-	-	-
PeopleOnStreet	-0.7	-0.3	0.0	-	-	-
Kimono	-0.3	0.0	0.0	-0.1	0.0	0.1
ParkScene	-0.1	-0.1	-0.1	0.0	0.0	0.0
Cactus	-0.4	0.2	0.2	-0.3	0.3	0.4
BasketballDrive	-0.6	-0.3	-0.1	-0.7	0.7	0.6
BQTerrace	-0.7	-0.2	0.1	-0.4	0.1	0.3
Class A average	-0.5	0.0	0.2	-	-	-
Class B average	-0.4	-0.1	0.0	-0.3	0.2	0.2
Average	-0.4	0.0	0.1	-0.3	0.2	0.3

Table 5. BD rate (in %) of all intra for MPM list based approach with diagonal only scan (no parsing issue)

	2x			1.5x		
	Y	U	V	Y	U	V
Traffic	-0.2	0.4	0.4	-	-	-
PeopleOnStreet	-0.4	0	0.2	-	-	-
Kimono	-0.2	0.1	0.1	-0.1	0.3	0.4
ParkScene	-0.1	0.1	0.1	-0.1	0.2	0.2
Cactus	-0.2	0.4	0.4	-0.3	0.4	0.4
BasketballDrive	-0.3	-0.2	0	-0.5	0.5	0.4
BQTerrace	-0.5	-0.1	0.2	-0.3	0.1	0.2
Class A average	-0.3	0.2	0.3	-	-	-
Class B average	-0.2	0.1	0.2	-0.2	0.3	0.3
Average	-0.3	0.1	0.2	-0.2	0.3	0.3

4. CONCLUSION

In this paper, we have presented a new coding algorithm for coding of intra prediction mode of enhancement layer in scalable extension of HEVC. We have also proposed a method to resolve the parsing dependency of enhancement layer on the base layer and illustrated that our proposal still shows gain with disabling MDCS.

5. REFERENCES

[1] G. Sullivan and J.R. Ohm, "Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC)," *ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, 2012.

[2] G.J. Han, J.R. Ohm, Woo-Jin Han, Woo-Jin Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1, 2012.

[3] C.A. Segall and G.J. Sullivan, "Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1121 –1135, sept. 2007.

[4] Yang Liu, G. Rath, and C. Guillemot, "Improved Intra Prediction for H.264/AVC Scalable Extension," in *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on*, oct. 2007, pp. 247 –250.

[5] J. Lainema, F. Bossen, W. Han, J. Min, and K. Ugur, "Intra Coding of the HEVC Standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1, 2012.

[6] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 620–636, 2003.

[7] J. Sole, R. Joshi, N. Nguyen, T. Ji, M. Karczewicz, G. Clare, F. Henry, and A. Duenas, "Transform Coefficient Coding in HEVC," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1, 2012.

[8] G. Bjontegaard, "Improvements of the BD-PSNR model," *ITU-T SG16 Q*, vol. 6, pp. 35, 2008.