Improved Lossless Compression Algorithm in HEVC

Kibaek Kim and Jechang Jeong
Department of Electronic and Computer Engineering
Hanyang University
Seoul, South Korea
e-mail: k2b0002@hanyang.ac.kr; jjeong@hanyang.ac.kr

***Abstract*—The state-of-the-art video-coding standard is High Efficiency Video Coding (HEVC), which achieves significant compression performance relative to H.264/AVC. HEVC is designed for lossy compression, however, and it is not ideal for lossless compression. To overcome the limitations of conventional coding in HEVC, we focus on the improvement of entropy coding for lossless compression. In this method, entropy coding is modified based on the characteristics of residuals in a lossless environment. The experimental results show that the bit savings of the proposed entropy coding is 1.46% on average when compared to HEVC lossless coding. Modified entropy coding with block-based intra prediction shows a bit savings of 7.99% on average. We observe that the combination of intra prediction and entropy coding is effective in screen-contents sequences.**

Keywords-HEVC; RDPCM; intra prediction; entropy coding; lossless compression.

I. INTRODUCTION

High Efficiency Video Coding (HEVC) is a state-of-the-art video-coding standard that outperforms other standards in several ways. HEVC aims to achieve significant compression improvements in the range of 50% output bit-rate reduction while still offering the same subjective quality as the previous video-coding standard, H.264/AVC (Advanced Video Coding). HEVC has been widely exploited for many uses, including in tablets, mobile phones, and high-definition televisions [1-3].

Lossy and lossless compression are two major classes of compression [4]. When compared with an original image, some data loss in the reconstructed image is allowed for in lossy compression. In lossless compression, in contrast, the reconstructed image is identical to the original image. Although HEVC supports both lossy and lossless compression, most of the coding techniques that are used in HEVC were designed for the improvement of coding performance in lossy compression. It is thus difficult to improve on the performance of lossless compression in the current HEVC structure.

Since lossless compression bypasses the transform, quantization, and in-loop filtering processes such as de-blocking filter and sample adaptive offset (SAO), the remaining processes are prediction and entropy coding, as shown in Fig. 1. The roles of prediction and entropy coding are therefore key for achieving higher compression. Intra prediction can be divided into block-based prediction and pixel-based prediction. Since the reconstructed image is identical to the original image, pixel-based prediction is generally more effective than block-based prediction, although pixel-based prediction may cause harmony-related problems with block-based structures in HEVC.

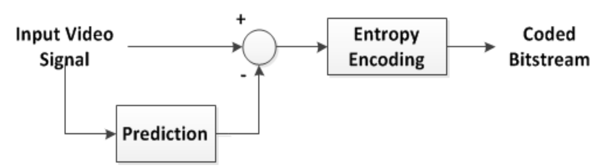


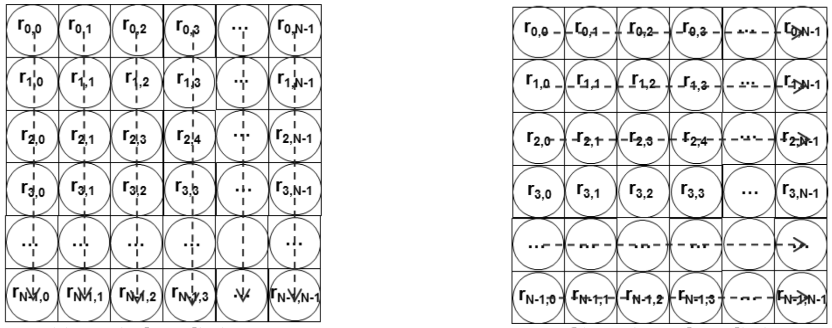
Figure 1. Flow chart of lossless compression

Entropy coding in HEVC focuses on the efficient coding of the coefficients that occur in lossy compression environments. The coefficients in lossless compression environments have different statistical characteristics than those in lossy compression environments. Considering the statistical characteristics of lossless compression, several researchers have proposed modifications related to entropy coding [5-7].

This paper focuses on the improvement of compression performance. For the prediction setup, a block-based scheme was considered as the intra prediction in this work. In the case of entropy coding, the level coding setup was modified based on the characteristic of the coefficient in a lossless compression environment. When intra prediction and entropy coding are combined to maximize compression performance, some compression improvement can be expected. In this paper, a block-based intra prediction method based on differential pulse-code modulation (DPCM) and several modifications of entropy coding have been combined and tested experimentally.

The remainder of this paper is organized as follows. Section 2 explains related work related to the proposed algorithm; Section 3 provides the proposed method for the entropy-coding process in a lossless coding environment. Section 4 includes an experimental evaluation, and Section 5 offers conclusions on the work.

 ISBN: 978-1-61208-490-9



(a) Vertical prediction **(b) Horizontal prediction**

Figure 2. Example of NxN RDPCM

II. RELATED WORK

A. Block-based Intra Prediction

In the case of block-based prediction, reference pixels that are located farther away than neighboring pixels may be used for prediction. In terms of prediction accuracy, block-based prediction may be reduced because of the block-based structure. In order to solve this problem, several studies on applying DPCM to the residual signals have been conducted; the output of these studies is based on "residual DPCM," or RDPCM [8-12]. While RDPCM is used to apply DPCM on the residuals to ensure pixel-based performance, it is also possible to maintain a block-based structure.

When intra prediction in HEVC is applied, the prediction pixels can be obtained from the reference pixels. We may then acquire differences between the original pixels and the prediction pixels, which are denoted as the residuals $r_{i,j}$ ($i,j \in \{0,1,\ldots N-1\}$).

As shown in Fig 2, RDPCM can be applied to the residuals $r_{i,j}$ that are obtained after intra prediction when the intra-prediction mode is vertical or horizontal. Suppose an NxN array R with elements $r_{i,j}$ in an array that includes the residual pixels obtained after intra prediction. When DPCM is applied to the NxN array R, we may attain the modified NxN array R' with elements $r'_{i,j}$ as a result of RDPCM.

The elements $r'_{i,j}$ obtained from the modified NxN array R' can be represented as follows when the intra-prediction mode is vertical:

$$r'_{i,j} = \begin{cases} r_{i,j}, & i = 0, 0 \le j \le (N-1) \\ r_{i,j} - r_{i-1,j}, & 0 < i \le (N-1), 0 \le j \le (N-1) \end{cases} \quad (1)$$

or as follows when the intra-prediction mode is horizontal:

$$r'_{i,j} = \begin{cases} r_{i,j}, & 0 \le i \le (N-1), j = 0 \\ r_{i,j} - r_{i,j-1}, & 0 \le i \le (N-1), 0 < j \le (N-1) \end{cases} \quad (2)$$

The elements $r'_{i,j}$ of R' are sent to the decoder instead of the elements $r_{i,j}$ of R. Through inverse RDPCM, the elements of $r_{i,j}$ of R can be reconstructed as follows, with the elements $r'_{i,j}$ of R' sent from the encoder when the intra-prediction mode is vertical:

$$r_{i,j} = \sum_{k=0}^{i} r'_{k,j}, 0 \le i, j \le (N-1) \quad (3)$$

or as follows when the intra-prediction mode is horizontal:

$$r_{i,j} = \sum_{k=0}^{j} r'_{i,k}, 0 \le i, j \le (N-1) \quad (4)$$

Algorithms related to RDPCM were proposed in [13-17]. To further improve prediction accuracy, the authors applied an additional prediction after the RDPCM process; they applied an additional prediction based on the spatial characteristic that the difference between neighboring pixels would be linearly increased or decreased according to the prediction direction in a block.

B. Entropy Coding of HEVC

After residuals are obtained from the prediction process, the residuals are then grouped into multiple square transform units (TUs). Though the basic unit of coefficient coding is TU, the coefficients are encoded per 4x4 sub-blocks. The syntax elements for the coefficient coding include *sig_coeff_flag*/*coeff_abs_level_g1_flag*/*coeff_abs_level_g2_flag*/*coeff_sign_flag*/*coeff_abs_level_remaining*. Each syntax element shows (respectively) whether or not (a) the coefficient is non-zero, (b) the absolute value of the coefficient is larger than one, and (c) the absolute value of the coefficient is larger than two; they also show (d) the sign information and (e) the remaining absolute-level value [18].

For the remaining absolute level of coefficient (except for 1 or 2), *coeff_abs_level_remaining* is binarized by a *k*-th order truncated Rice and a (*k*+1)th order Exp-Golomb (EG*k*). HEVC supports a range from 0 to 4 for the Rice parameter *k*, as shown in Table 1. When an absolute level of coefficient is greater than the threshold, EG*k* is added. The initial value of the Rice parameter *k* is zero. During the coding process, the Rice parameter depends on the current parameter and the previously coded value of *coeff_abs_level_remaining*.

TABLE I. BINARIZATION TABLES ACCORDING TO *K*-TH ORDER TRUNCATED RICE CODE

v	Codeword				
	k = 0	*k* = 1	*k* = 2	*k* = 3	*k* = 4
0	0	0.0	0.00	0.000	0.0000
1	10	0.1	0.01	0.001	0.0001
2	110	10.0	0.10	0.010	0.0010
3	1110	10.1	0.11	0.011	0.0011
4	11110	110.0	10.00	0.100	0.0100
5	111110	110.1	10.01	0.101	0.0101
6	1111110	1110.0	10.10	0.110	0.0110
7	11111110	1110.1	10.11	0.111	0.0111
8	111111110	11110.0	110.00	10.000	0.1000
9	1111111110	11110.1	110.01	10.001	0.1001
10	11111111110	111110.0	110.10	10.010	0.1010
11	111111111110	111110.1	110.11	10.011	0.1011
12	1111111111110	1111110.0	1110.00	10.100	0.1100
13	11111111111110	1111110.1	1110.01	10.101	0.1101

The updated Rice parameter *k* is always greater than the previous *k*. The increment of the parameter is determined as per the following condition:

$$If\ |x| > 3\cdot 2^k,\ then\ k' = \min(k+1, 4) \quad (5)$$

where k' denotes the updated Rice parameter and $|x|$ is an absolute value of the current coefficient [18].

III. PROPOSED METHOD FOR LOSSLESS CODING

In lossy compression, the magnitudes of the scanning position with a low index are relatively high compared to the magnitude of the scanning position with a high index [18][19]. The magnitude of each scanning position is similar, however, since lossless compression bypasses the transform and quantization processes. This means that the magnitudes of each coefficient in lossless compression are independent of the scanning position; it has also been observed that the range of non-zero coefficients in lossless compression is generally wide in comparison to lossy compression [6]. This insight allows us to know that the current design of the entropy coding in HEVC does not work efficiently in lossless compression environments.

Several algorithms have been proposed to achieve efficient coding in lossless compression [5-7]; these studies include (respectively) changing the scan order according to prediction mode, extending the binarization tables for level coding, and changing the rule of binarization-table selection. A binarization table related to these algorithms is shown in Table 1.

We propose a modified binarization-table-selection method based on the characteristics of residuals in a lossless compression environment; this method is based on decision the updated Rice parameter [6].

In HEVC, an absolute value of current coefficient is used for renewal of Rice parameter. In [6], the previous coefficients in addition to current coefficient are used for the renewal. The renewal is based on the weighting sum of the previous encoded coefficients. In calculation of weighting sum, the previous zero coefficients may influence according to distribution of non-zero coefficients. In the proposed method, only the previous non-zero coefficients of current coefficient position influence the prediction of the next non-zero coefficient and the binarization-table selection. This point is different from that proposed in [6]. We explain the details of the proposed method in the following sub-section. We also briefly describe other modifications.

A. Proposed Entropy Coding

Absolute levels of previous non-zero coefficients in addition to the current coefficient are used to calculate $T(s_i)$ in the proposed method. Since $T(s_i)$ is considered for the prediction of the next non-zero coefficient, zero coefficients may have a negative influence on the prediction. In the proposed method, the absolute levels of the previous non-zero coefficients are considered in the calculation of $T(s_i)$.

The decision procedure of $T(s_i)$ for determining the level-binarization table is described as follows:

$$T(s_i) = \frac{1}{w_i}\sum_{k=p}^{i} s_k \qquad (6)$$

$$w_i = \begin{cases} b\ , & if\ nz_i = 1 \\ b+1\ , & if\ nz_i = 2 \\ b+2\ , & if\ nz_i = 3 \\ b+3\ , & otherwise \end{cases} \qquad (7)$$

In Equation (6), *i* and *p* denote the current scanning position and the past non-zero scanning position, respectively. In Equation (7), nz_i is the number of non-zero coefficients that contain the current coefficient in a 4x4 sub-block. According to nz_i, the sharing value w_i is then determined; in addition, *p* is influenced by w_i. The (w_i-1)-th previous non-zero coefficient is the starting position *p* for the summation. The parameter *b* for influencing w_i reflects the sharing value and is determined empirically (b = 1).

For the renewal of Rice parameter, the calculated T is used as the input value instead of $|x|$ in (5). According to the result, Rice parameter is determined for encoding of the absolute level of the next non-zero coefficient. In contrast with HEVC, the Rice parameter *k* for the binarization-table selection can increase or decrease according to the decision procedure.

B. Other Modifications of Entropy Coding

TABLE II. MODIFIED SCANNING PATTERN ACCORDING TO MODE

Mode	Block size			
	4x4	*8x8*	*16x16*	*32x32*
Mode (H)	Diag	V->H	H->V	Diag
Mode (V)	Diag	H->V	V->H	Diag
Remainder	Diag	Diag	Diag	Diag

Three scanning patterns are supported in HEVC: (1) diagonal upright, (2) horizontal, and (3) vertical [18]. The scanning pattern is adaptively determined according to intra mode. In lossless compression, a correlation in the prediction direction remains. As shown in Table 2, the adaptive-scanning patterns are thus changed in our method [5]. Considering the characteristic that some coefficients with a very high magnitude will appear, the range of the Rice parameter k is extended from 0 to 6 [7].

IV. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed method, we implemented the method using HM 10.0 reference software; the experimental condition was "Intra, Main," which is one of the most common test conditions, discussed in [20][21].

In order to apply a lossless coding mode, the quantization parameter (QP) was set to 0, while the TransquantBypassFlag, CUTransquantBypassFlagValue, and LosslessCUEnabled parameters were set to 1. The transform and quantization processes were skipped. Other details of the experimental conditions are shown in Table 3. Various HEVC test sequences were used for the simulations. In Table 3, CU and LCU mean coding unit and largest coding unit, respectively.

TABLE III. EXPERIMENTAL CONDITION

Parameter	Value	Description
CUWidth	64	LCU size (64x64)
CUHeight	64	
IntraPeriod	1	All intra coding
QP	0	Lossless coding
Reference Sample Smoothing	off	
Block Boundary Filtering	off	
LoopFilterDisable	1	No in-loop filtering
SAO	0	

The compression performance was evaluated for bit savings as follows. In bit savings, positive values denote an improvement in coding efficiency.

$$Bit\ saving(\%) = \frac{Bitrate_{HEVC} - Bitrate_{proposed}}{Bitrate_{HEVC}} \times 100 \qquad (8)$$

Several techniques for lossless compression were either evaluated independently, or a combination of tools was evaluated. In terms of intra prediction, the block-based method [17] in addition to RDPCM was exploited in our technique. Three components of entropy coding from section 3 were combined as the proposed entropy coding. In order to evaluate the compression performance, five categories were partitioned: proposed entropy coding (M1), RDPCM (M2), block-based prediction [17] (M3), a combination of A and B (M4), and a combination of A and C (M5). Table 4 summarizes the experimental results of the bit-savings tests. The table shows the results of M1, M2, M3, M4, and M5, which are compared with HM 10.0 lossless intra coding.

As shown in Table 4, the bit savings of M1 were 1.46% on average in total test sequences. Especially, the test did show meaningful results in class F. Class F shows that the bit savings were 3.48% on average. The sequences in that class

 ISBN: 978-1-61208-490-9

consisted of screen contents. Compared to natural images, screen-content images tend to experience many changes between adjacent pixels. Some coefficients therefore may have a very high magnitude. Since the proposed entropy-coding method considers the characteristics in a lossless compression environment, the method can prevent wasteful bits from representing non-zero coefficients that have a very high magnitude.

In the case of block-based prediction, the bit savings of M2 and M3 were 5.83% and 7.29%, respectively. Since intra prediction is comprised of pixel-by-pixel DPCM, intra prediction can provide better coding efficiency compared to HM10.0 lossless coding. M3 achieved a small improvement over M2, since M3 applies additional prediction after the RDPCM process.

In order to optimize lossless compression, we experimented with a combination of intra prediction and entropy coding. As shown in Table 4, the bit savings of combinations M4 and M5 were 6.43% and 7.99% on average, respectively. When compared with M2 and M3, test results showed additional coding gains of 0.6% and 0.7% (respectively) on average in total sequences. In addition, M4 and M5 showed significant improvement in class F of screen content. The difference of bit savings compared to M2 and M3 were 2.13% and 2.11%, respectively.

TABLE IV. EXPERIMENTAL RESULTS OF BIT SAVINGS

Bit savings (%)						
Class	Sequence	M1	M2	M3	M4	M5
A	Traffic	1.58	8.20	11.34	8.88	12.22
	PeopleOnStreet	1.74	7.57	11.85	8.41	12.98
	Nebuta	-0.19	5.65	10.12	4.79	9.64
	SteamLocomotive	1.21	6.04	11.39	6.78	12.53
Average of class A		1.09	6.87	11.18	7.22	11.84
B	Kimono	1.67	4.28	6.30	5.17	7.36
	ParkScene	0.87	5.01	7.55	5.26	7.94
	Cactus	0.38	2.17	2.73	1.99	2.56
	BQTerrace	-0.02	2.67	3.32	2.04	2.66
	BasketballDrive	0.67	2.49	2.80	2.64	2.99
Average of class B		0.71	3.33	4.54	3.42	4.70
C	RaceHorses	1.15	5.17	7.51	5.53	8.09
	BQMall	1.33	4.47	4.86	4.93	5.27
	PartyScene	0.54	3.60	3.89	3.31	3.59
	BasketballDrill	1.50	1.79	2.07	2.71	2.98
Average of class C		1.13	3.76	4.58	4.12	4.98
D	RaceHorses	0.97	6.06	8.88	6.02	9.09
	BQSquare	1.15	2.68	2.62	3.19	3.15
	BlowingBubbles	0.45	3.82	4.01	3.28	3.45
	BasketballPass	1.86	8.98	10.71	9.69	11.54
Average of class D		1.11	5.38	6.55	5.55	6.81
E	FourPeople	1.67	8.75	10.87	9.49	11.77
	Johnny	1.29	7.59	8.53	8.22	9.26
	KristenAndSara	1.35	8.48	9.69	9.15	10.44
Average of class E		1.43	8.28	9.70	8.95	10.49
F	BasketballDrillText	1.91	2.56	2.69	3.70	3.77
	ChinaSpeed	2.83	11.54	10.23	13.41	11.91
	SlideEditing	6.00	7.97	6.05	12.14	10.43
	SlideShow	3.18	12.35	14.96	13.69	16.24
Average of class F		3.48	8.61	8.48	10.74	10.59
Total average		1.46	5.83	7.29	6.43	7.99

TABLE V. COMPARISON OF CODING TIMES

Overall	M2	M3	M4	M5
Encoder	99%	99%	103%	104%
Decoder	92%	89%	95%	93%

Table 5 describes the encoding time and decoding time of each method when each was compared with HM 10.0 lossless intra coding. In the case of block-based prediction, there was no increase in coding time, since high prediction accuracy leads to the reduction of residual data to be encoded or decoded in our method. The proposed entropy coding, when combined with block-based prediction method, led to increases in coding time because of the additional processes we use related to entropy coding.

V. CONCLUSION

In order to overcome the limitations of the conventional coding structure based on lossy coding, we have proposed a method for improving performance in lossless compression environments. In this method, entropy coding was modified based on the characteristic of residuals in a lossless environment. In comparison with HEVC, the experimental results showed that the proposed method achieved a bit reduction of 1.46% on average. Modified entropy coding with block-based intra prediction achieved bit reductions of 7.99% on average. In particular, the combination of intra prediction and entropy coding showed good compression performance in screen-contents sequences.

There is still a need to improve on the coding gains in some of the sequences, however. Further study is thus required to improve coding performance in lossless coding environments.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(ITP) grant funded by the Korea government(MSIP) (No.2016-B0101-15-1377, Development of 16-channel Multi-viewer, H/W Encoder, and Transcoder Systems for 4K UHD Broadcasting)

REFERENCES

[1] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, Dec. 2012, pp. 1649-1668.

[2] B. Bross et al, “High Efficiency Video Coding (HEVC) text specification draft 10,” JCTVC-L1003, Geneva, Switzerland, Jan. 2013.

[3] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan and T. Wiegand, “Comparison of the coding efficiency of video coding standards including High Efficiency Video Coding (HEVC),” IEEE Trans. Curcuits Syst. Video Technol., vol. 22, no. 12, Dec. 2012, pp.1649-1668.

[4] K. Sayood, Lossless Compression Handbook, San Diego, Academic Press, 2003.

[5] W. Gao, M. Jiang, and H. Yu, “On lossless coding for HEVC,” in: Proc. SPIE, Feb. 2013, pp. 8666–8674.

[6] J. Choi and Y. Ho, “Improved residual data coding for high efficiency video coding lossless extension,” in Proc. RIVF, 2013, pp. 18-21.

[7] J. Choi and Y. Ho, “Differential pixel value coding for HEVC lossless compression,” Advanced video coding for next-generation multimedia services, INTECH; 2013.

[8] Y. Lee and K. Han, “Lossless coding for professional extensions,” JVT-L017, Redmond, USA, Jul. 2004.

[9] Y. Lee, K. Han, and S. Lim, “Lossless intra coding for improved 4: 4: 4 coding in H. 264/MPEG-4 AVC,” Geneva, Switzerland, JVT-P016, Geneva, Switzerland, Jul. 2005.

[10] Y. Lee and K. Han, “Complexity of the proposed lossless intra,” JVT-Q035r1, Nice, France, Oct. 2005.

[11] Y. Lee, K. Han, and G. J. Sullivan, “Improved lossless intra coding for H. 264/MPEG-4 AVC,” IEEE Trans. Image Process., vol. 15, no. 9, Sep. 2006, pp. 2610-2615.

[12] S. Lee, I. Kim, and C. Kim, “AHG7: Residual DPCM for HEVC lossless coding,” JCTVC-L0117, Geneva, Switzerland, Jan. 2013.

[13] J. Kwak and Y. Lee, “Secondary residual transform for lossless intra coding in HEVC,” J. of Broadcast Eng., vol. 17, no. 5, Sep. 2012, pp. 734-741.

[14] S. Hong and Y. Lee, “CR-DPCM for lossless intra prediction method in HEVC,” J. of Broadcast Eng., vol. 19, no. 3, Sep. 2014, pp. 307-315.

[15] S. Hong, J. Kwak, and Y. Lee, “Cross residual transform for lossless intra-coding for HEVC,” Signal Process. Image Commun., vol. 28, no. 10, Nov. 2013, pp. 1335-1341.

[16] Y. Lee and S. Hong, “Cross residual DPCM for HEVC lossless coding,” JCTVC-N0113, Vienna, Austria, Aug. 2013.

[17] K. Kim, G. Jeon, and J. Jeong, “Improvement of implicit residual DPCM for HEVC,” in Proc. SITIS, 2014, pp. 652-658.

[18] J. Sole et al, “Transform coefficient coding in HEVC,” IEEE Trans. Circuits syst. video technol., vol. 22, no. 12, Dec. 2012, pp. 1765-1777.

[19] D. Marpe, H. Schwarz, and T. Wiegand, “Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 13. no. 7, Jul. 2003, pp. 620-636.

[20] HEVC Test Model. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVC Software/tags/HM-10.0/.

[21] F. Bossen, “Common HM test conditions and software reference con figurations,” JCTVC-K1100, Shanghai, China, Oct. 2012.