LOW BIT-RATE INTRA CODING SCHEME BASED ON CONSTRAINED QUANTIZATION AND MEDIAN-TYPE FILTER

Chen Chen and Bing Zeng

Department of Electronic & Computer Engineering
The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
{cchenaf,eezeng}@ust.hk

ABSTRACT

This paper presents a new intra coding scheme for low bit-rate video compression. We first propose an improved codec architecture based on HEVC encoder. Then we divide pixels in each prediction block into two parts: three quarters of pixels are coded via a smart padding technique together with a constrained quantization algorithm (leading to a significantly improved quality); whereas the other quarter are reconstructed according to median-type filtering by utilizing the 8-neighboring reference samples after all blocks have been encoded and reconstructed. Experimental results show that about 3% BD-rate reduction has been achieved both for luma and chroma components without apparent increase of encoding complexity with respect to the original HEVC intra coding.

***Index Terms*—** low bit-rate, video coding, smart padding, constrained quantization, median-type filter

1. INTRODUCTION

High Efficiency Video Coding (HEVC) is a newly established video coding standard in Jan. 2013 as the successor of H.264/AVC by joint collaborative team on video coding (JCT-VC), whose primary goal is to achieve 50% bit-rate reduction as compared with H.264/AVC under the same perceptual quality [1]. More and more people are inspired to dedicated in research on low bit-rate video coding with the rapid growth of demand for high-definition videos with 1080p or even 4K resolutions and the limited storage space or network bandwidth, especially pursing intra coding techniques since there are still huge space for I-frame compression.

While inheriting most features of H.264/AVC, HEVC introduces a large number of new and advanced coding techniques in its intra-frame coding. Unlike H.264/AVC that fixes basic coding units at 16×16 macro-blocks, HEVC enables a so-called coding tree unit (CTU) with size from 16×16 up to 64×64, which can be further split down to 8×8 coding unit (CU) determined by a coded frame's texture complexity. The quad-tree structure of HEVC allows the prediction unit (PU) size varying from 4×4 to 64×64 and transform unit (TU) size from 4×4 to 32×32 flexibly according to textures in a frame [2]. The large scope of CU size and self-adaptive PU and TU size selection provide HEVC intra coding a higher compression efficiency as compared to H.264/AVC.

In order to achieve better intra prediction, HEVC expands H.264/AVC's 9 intra prediction modes to as many as 35, including 33 angular modes, the DC mode, and a planar mode for all block sizes on luma coding. Especially, the introduction of planar mode, defined as an average of two linear predictions, aims to replicate the benefits of the plane mode while preserving continuities along block boundaries [2]. In terms of chroma, HEVC enables 5 modes, including DC, planar, vertical, horizontal, and the mode used by corresponding luma block. Thus, HEVC can do intra-prediction more accurately, leading to a smaller residual energy, while spending more bits on intra modes representation at the same time as compared to H.264/AVC. Furthermore, 4×4 DST [3] is newly introduced to apply on 4×4 Luma block and adaptive coefficient scanning order (diagonal, vertical and horizontal) are utilized to replace traditional zigzag scanning [4].

Although HEVC performs intra prediction more precisely, the residual energy could still be large, leading to a low compression efficiency for I frames. Therefore, people have been dedicated to looking for breakthroughs on intra coding. In [5], new chroma intra prediction modes based on a linear model have been proposed, which can improve the chroma intra coding efficiency by about 6% BD-rate [6] reduction but no noticeable gain for luma.

As we all know, intra prediction makes full use of the spatial correlations among neighboring pixels by performing predictions along different directions based on the reconstructed reference pixels (that are above and to the left of the current block). After performing transform on the residuals, both H.264/AVC and HEVC quantize transform coefficients in a transform block (TB) using pre-defined quantization parameters (QP). These quantization errors lead to coding errors on all pixels within the block (obtained after the inverse transform). A theory has been put forward in [7] that the predicted residuals still usually have strongest correlation in the direction of the prediction so that the energy of the residuals

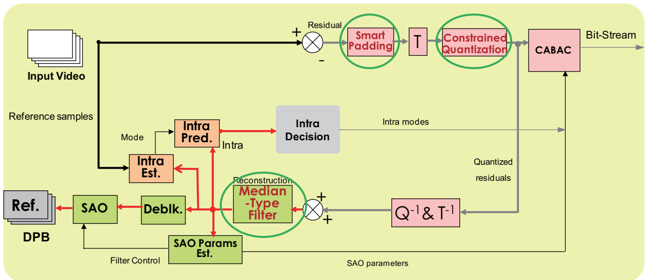
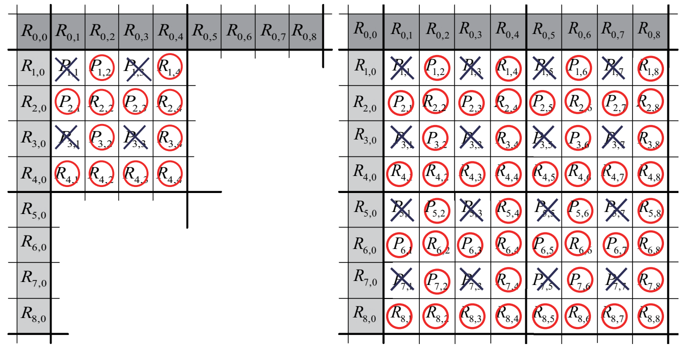


Fig. 1. Improved encoder architecture based on HEVC intra encoding structure, including three new introduced modules—Smart Padding, Constrained Quantization and Median-Type filter in green ellipses.



(a) 4×4 TB (b) 8×8 TB

Fig. 2. Coding patterns used in our constrained quantization for 4×4 and 8×8 TBs (positions with red circles are coded by constrained quantization method and positions with blue crosses are reconstructed by median-type filters).

would become smaller when they are located closer to the reference pixels. Consequently, Zhu et al. [8] propose a so-called constrained quantization technique that focus on coding a selected portion of pixels in a block while ignoring the others.

Zeng et al. [9] demonstrate that median-type filters are typically more effective than linear interpolators when enhancing image resolution by interpolation, especially for low-quality images. Therefore, in this paper, we propose an intra coding scheme based on constrained quantization and median-type filters for low bit-rate video coding.

The rest of the paper is organized as follows. In Section 2, we detailed describe the new intra coding method, including the use of a smart padding technique, the constrained quantization algorithm and median-type filter. Some experimental results are given in Section 3. Finally, some conclusions and discussions are presented in Section 4.

2. PROPOSED INTRA CODING SCHEME

2.1. Improved Encoder Architecture

An intelligent quantization algorithm has been introduced in [8]. Such a constrained quantization focuses on coding of a selected subset of pixels (within each block) and is able to produce a remarkably improved quality (on all selected pixels) without any increasing of bit-rate. The principle of this quantization is to shape the distortion of all selected pixels within each N×N block onto the other pixels in the same block according to a precisely-determined adjustment on transform coefficients. Experiments in [8] demonstrate that such quantization algorithm can achieve about 0.5dB and even up to 1dB gain on average under the same bit-rate when considering coding only three quarters of pixels in each block.

Inspired by the constrained quantization approach and based on the original HEVC intra codec structure, Fig. 1 detailed shows the improved encoder architecture, including three major changes—the smart padding technique after obtaining residuals, constrained quantization instead of traditional quantization and newly proposed median-type filter applied after all blocks have been encoded and reconstructed. Each module will be discussed in the following subsections.

2.2. Pattern Design and Smart Padding

In this article, we propose to encoding three quarters of pixels in a TB and completely employ newly proposed method to substitute HEVC's original intra modes, aiming at significantly reducing algorithm complexity and further enhancing compression efficiency for low bit-rate video coding.

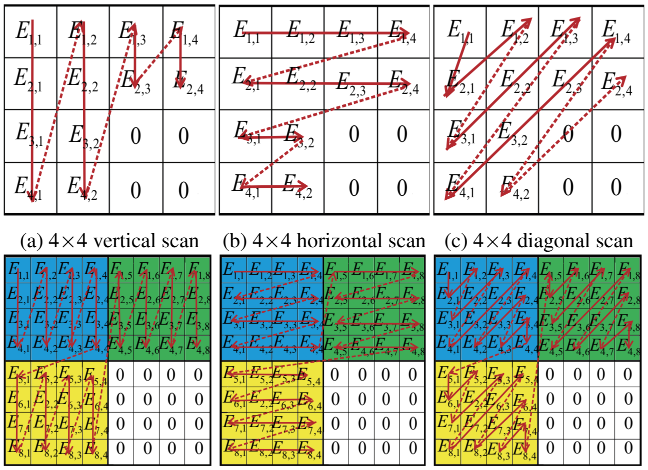
In order to decrease the difficulty of reconstructing missing pixels and the complexity of encoder, we propose to adopt the pattern for each 4×4 TB as shown in Fig. 2(a) since all 8 neighbourhoods of missing pixels are already known after coding and it can be expanded in a straightforward way to the 8×8 pattern as shown in Fig. 2(b). Expansions to 16×16 and 32×32 can also be made accordingly. In Fig. 2, the pixels marked by red circles (classified into $\mathbb{Q}$) are coded by the constrained quantization based approach and the pixels marked by blue crosses (classified into $\mathbb{F}$) are reconstructed by median-type filters after all blocks are encoded.

Since only three quarters of residuals in a TB should be taken into consideration, we propose to do a padding on locations $(k, l) \in \mathbb{F}$ such that, in the transform domain, we can force a quarter of transform coefficients to be 0 - this is exactly the smart padding technique developed in [10].

Define $\mathbf{H}_{m\times m}$ as the unitary transform matrix, $\boldsymbol{\varepsilon}$ as the prediction error matrix, and $\mathbf{E}$ as the transform coefficient matrix, respectively. Based on the transform equation $\mathbf{E} = \mathbf{H} \cdot \boldsymbol{\varepsilon} \cdot \mathbf{H}^T$, we can formulate $m^2/4$ equations

$$\sum_{q=1}^{m} \left\{ h_{y_s,q} \cdot \left[\sum_{p=1}^{m} \left(h_{x_s,p} \cdot e_{p,q} \right) \right] \right\} = 0 \qquad (1)$$

where $h_{y_s,q}$ is the $(y_s, q)^{th}$ element of $\mathbf{H}$, $e_{p,q}$ is the $(p, q)^{th}$ element of $\boldsymbol{\varepsilon}$, and $m^2/4$ pairs $\{(x_s, y_s)|s = 1, 2, \cdots, m^2/4\}$



(a) 4×4 vertical scan (b) 4×4 horizontal scan (c) 4×4 diagonal scan

(d) 8×8 vertical scan (e) 8×8 horizontal scan (f) 8×8 diagonal scan

Fig. 3. Direction-adaptive scanning orders of transform coefficients for 4×4 and 8×8 TBs. Locations whose coefficient are 0 will be used to derive smart padding matrix.

are the selected quarter positions where the transform coefficients are 0 (4×4 and 8×8 smart padding pattern are shown in Fig. 3. 16×16 and 32×32 pattern can be naturally derived accordingly). Finally, the prediction errors located at position $(k, l) \in \mathbb{N}$ are smart padded according to the prediction errors located at position $(i, j) \in \mathbb{Q}$ via

$$\begin{bmatrix} e'_{k_1,l_1} \\ e'_{k_2,l_2} \\ \vdots \\ e'_{k_{m^2/4},l_{m^2/4}} \end{bmatrix} = \mathbf{\Psi}_{m^2/4\times 3m^2/4} \begin{bmatrix} e_{i_1,j_1} \\ e_{i_2,j_2} \\ \vdots \\ e_{i_{3m^2/4},j_{3m^2/4}} \end{bmatrix} \tag{2}$$

where $\mathbf{\Psi}_{m^2/4\times 3m^2/4}$ is a matrix derived from Eq. (1)

2.3. Constrained Quantization

In our case, we need to handle three quarters of pixels in a block. First, we need to derive a matrix $\mathbf{B} = [\mathbf{B}_0, \mathbf{B}_1]$ to be used in the constrained quantization method [8]. $\mathbf{B}$ is an $3m^2/4 \times m^2$ matrix and can be derived from the transform matrix $\mathbf{H}$ and the coefficients scanning order; whereas $\mathbf{B}_0$ is an $3m^2/4 \times 3m^2/4$ upper triangular matrix.

Since $\mathbf{\Psi}_{m^2/4\times 3m^2/4}$ is well-conditioned as described in Section 2.2 and the constrained quantization method requires that the diagonal elements of matrix $\mathbf{B}_0$ be large enough (e.g., 0.5) for the purpose of keeping the transform coefficients to be small in magnitude [8], we need to tune carefully the transform coefficients scanning order based on the HEVC's scheme. After different tries, we found that the scanning orders as shown in Fig. 3 serve the purpose well.

For the reason that HEVC introduces three kinds of coefficients scanning order for 4×4 and 8×8 TBs [11], Fig. 3 illustrates in detail the alternative scanning order (vertical, horizontal, and diagonal) for the constrained quantization coding method. When the TB size is beyond 8×8 (namely 16×16 and 32×32), only diagonal scan is allowed, so that the diagonal scan pattern for larger blocks can be derived from Fig. 3(c) and (f) based on 4×4 sub-block scanning order [12].

Combining the smart padding with the alternative scanning orders, the last quarter of transform coefficients along scanning order are exactly 0. Thus, only the first three quarters of transform coefficients need to be scanned and transmitted. Total energy is now compacted into the first three quarters of coefficients, thus significantly reducing the bit-count for the entropy coding.

The iterative algorithm for this constrained quantization process is summarized as follows:

Loop: *For* $i = 3m^2/4 : -1 : 1$, *following inverse coefficient scanning order, calculate the compensation amount*

$$\Delta_i = \frac{1}{b_{i,i}} \sum_{j=i+1}^{m^2} b_{i,j}(E_j - \hat{E}_j) \tag{3}$$

and perform the normal quantization on the compensated E_i

$$\hat{E}_i = round\left(\frac{E_i + \Delta_i}{qp}\right) \tag{4}$$

In these two equations, $b_{i,j}$ is the $(i, j)^{th}$ element of matrix $\mathbf{B}$ and $\hat{E}_i$ is the constrained quantized transform coefficient.

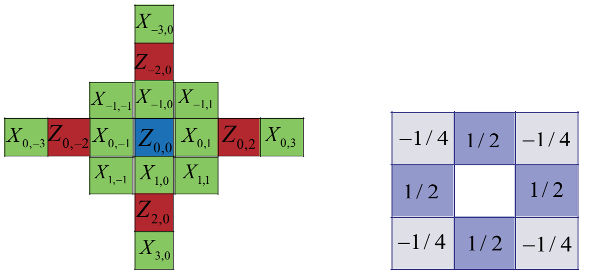
Notice that such coefficients adjusting aims at transferring distortion of pixels located in $\mathbb{Q}$ to those located in $\mathbb{F}$, thus further enhancing the rate-distortion performance of pixels located in $\mathbb{Q}$.

2.4. Median-Type Filter

After smart padding and constrained quantization based coding for all blocks, three quarters of the whole pixels (located in $\mathbb{Q}$) in a frame have been reconstructed. The remaining quarter of pixels (located in $\mathbb{F}$) need to be recovered by some interpolators or filters. Zeng, et al. [9] investigate the performance of finite impulse response (FIR)-median hybrid filters and conduct some experiments that reveal such median-based interpolators outperform significantly as compared with linear schemes in many cases. Hence we adopt median-type filters as our reconstruction interpolators.

Observing the coding pattern displayed in Fig. 2, all the 8 neighborhoods of locations in $\mathbb{Q}$ have been reconstructed, leading a great advantage for us to apply median-type filters. Fig 4(a) illustrates the median-type filtering structure for reconstructing the central pixel $Z_{0,0}$. We employ a 7-points median-type filter consisting of 4 neighborhoods, a vertical FIR filter, a horizontal FIR filter and an 8-neighboring FIR filter as:

$$Z_{0,0} = Median\{X_{0,-1}, X_{0,1}, X_{-1,0}, X_{1,0}, X_H, X_V, X_{FIR}\} \tag{5}$$



(a) Median-type filtering structure (b) Parameters of FIR filter

Fig. 4. A sketch showing median-type filtering structure

where X_H,X_V are pre-designed FIR filters defined as:

$$X_H = [-1/2, 1, 1, 1/2] \cdot [X_{0,-3}, X_{0,-1}, X_{0,1}, X_{0,3}]^{\mathrm{T}} \quad (6)$$

$$X_V = [-1/2, 1, 1, 1/2] \cdot [X_{-3,0}, X_{-1,0}, X_{1,0}, X_{3,0}]^{\mathrm{T}} \quad (7)$$

and X_{FIR} is the 8-neighboring FIR filter with parameters listed in Fig. 4(b) and calculated by $X_{FIR} = \vec{X} \cdot \vec{P}$ where

$$\vec{X} = [X_{-1,-1}, X_{-1,0}, X_{-1,1}, X_{0,-1}, X_{0,1}, X_{1,-1}, X_{1,0}, X_{1,1}] \quad (8)$$

$$\vec{P} = [-1/4, 1/2, -1/4, 1/2, 1/2, -1/4, 1/2, -1/4]^{\mathrm{T}} \quad (9)$$

Finally, after applying the three steps described in above subsections, a whole frame has been encoded and reconstructed.

3. EXPERIMENT RESULTS

This section provides some experimental results for the newly proposed median-type filtering-based coding method and compares it against the original HEVC intra coding. Simulations are mainly conducted on a range of HEVC standard test video sequences with high resolutions. Performance of the proposed algorithm is compared against the algorithm implemented in the state-of-the-art HM16.6 [13] reference software using the default *encoder_intra_main* configuration following HM common test conditions (CTC) [14]. Furthermore, we take the average over 60 I frames to ensure the correctness of our experimental results.

The test set contains 16 sequences, split into four classes, depending on their resolutions: class A (4K), class B (1080p), class C (WVGA) and class E (720p). We sample 4 points with QP = {37, 40, 43, 46} for each low bit-rate rate-distortion (R-D) performance experiment. The coding efficiency results are presented in Table 1 as the percentage of bit-rate savings (BD-rate [6]) with respect to the HM-16.6 main profile anchor. Statistics reveal that our algorithm obtain about 3% performance gain on average and can be up to 5.1% for the sequence "SteamLocomotive" as compared with the original HEVC's intra coding scheme. Especially for 4K sequences, the coding gain is much better than that of WVGA sequences. In terms of the algorithm complexity, the encoding time increases by 10% on average and the decoding time increases by average 70% because of introduction of the median-type filters which need sorting operation.

Table 1. BD-rate reduction under different sequences compared with HM16.6 (60 I frames, *encoder_intra_main* configuration)

Sequence	QP 37-46		
	Y BD-rate	U BD-rate	V BD-rate
Traffic	-4.05%	-3.11%	-4.65%
PeopleOnStreet	-3.84%	-2.80%	-4.10%
Nebuta	-4.52%	-5.66%	-5.93%
SteamLocomotive	-5.14%	-6.26%	-2.57%
Avg	**-4.39%**	**-4.46%**	**-4.31%**
Kimono	-4.40%	-5.26%	-5.11%
ParkScene	-3.95%	-3.91%	-2.20%
Cactus	-3.18%	-4.25%	-2.78%
BasketballDrive	-2.78%	-3.20%	-2.78%
BQTerrace	-1.30%	-4.67%	-2.38%
Avg	**-3.12%**	**-4.26%**	**-3.05%**
BasketballDrill	-1.72%	-2.58%	-1.22%
BQMall	-0.62%	-2.10%	-1.63%
PartyScene	+4.05%	+1.86%	+1.92%
RaceHorses	-4.70%	-1.78%	-1.08%
Avg	**-0.75%**	**-1.15%**	**-0.50%**
FourPeople	-2.43%	-3.26%	-2.77%
Johnny	-2.71%	-3.45%	-5.76%
KristenAndSara	-1.63%	-2.46%	-3.00%
Avg	**-2.26%**	**-3.06%**	**-3.84%**
Avg BD-rate	**-2.68%**	**-3.41%**	**-2.87%**
Encoding time	**110%**		
Decoding time	**170%**		

4. CONCLUSION

In this paper, we have proposed a new intra coding scheme based on smart padding, constrained quantization and median-type filters. Our algorithm codes only three quarters of pixels in each TB, while the remaining quarter in the same TB are reconstructed by median-type filtering utilizing the spatial correlation of neighboring pixels. Experimental results show that our proposed algorithm achieves about 2.7% BD-rate reduction for luma component and more than 3% compression gain for chroma components on average with respect to the original HEVC intra coding. The algorithm complexity for encoder basically remains unchanged but the decoding time increases by 70%.

Compared with the results of HEVC intra coding, compression efficiency has been further improved for low bit-rate cases and encoding time has no noticeable increasing. But due to the limitation of interpolation or filtering, there is no performance gain for high bit-rate circumstance. In addition, there is still some lose for specific sequences with complicated textures such as "PartyScene". So our future work will focus on improving robustness of the algorithm, the R-D performance for high bit-rate cases and further reducing decoding complexity.

5. REFERENCES

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

[2] J. Lainema, F. Bossen, W.-J. Han, J. Min and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2012.

[3] A. Saxena and F. C. Fernandes, "DCT/DST-Based transform coding for intra prediction in image/video Coding," *IEEE Trans. Image Processing*, vol. 22, no. 10, pp. 3974–3981, Oct. 2013.

[4] T. Wiegand, G. J. Sullivan, G. Bjøntegaard and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[5] X. Zhang, O. C. Au, J. Dai, C. Pang and F. Zou, "New chroma intra prediction modes based on linear model for HEVC," *IEEE International Conference on Image Processing, ICIP 2012*, Orlando, Florida, USA, Sep. 30–Oct. 3, 2012.

[6] G. Bjøntegarrd, *Calculation of average PSNR differences between RD curves*, VCEG-M33, 13th ITU-T Video Coding Experts Group (VCEG) Meeting, Austin, Texas, USA, Apr. 2001.

[7] M. Karczewicz, *Improved intra coding*, VCEG-AF15, 32nd ITU-T Video Coding Experts Group (VCEG) Meeting, San Jose, CA, USA, Apr. 2007.

[8] S. Zhu and B. Zeng, "Constrained quantization in the transform domain with applications in arbitrarily-shaped object coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1385–1393, Nov. 2011.

[9] B. Zeng and A. N. Venetsanopoulos, "Image interpolation based on median-type filters," *Optical Engineering*, vol. 37, no. 9, pp. 2472-2482, Sep. 1998.

[10] G. Shen, B. Zeng, and M. L. Liou, "Arbitrarily-shaped transform coding based on a new padding technique," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 1, pp. 67–79, Jan. 2001.

[11] J. Sole, R. Joshi, N. Nguyen, T. Ji, M. Karczewicz, G. Clare, F. Henry and A. Dueñas, "Transform coefficient coding in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1765–1777, Dec. 2012.

[12] J. Sole, R. Joshi, and M. Karczewicz, *Non-CE11: Diagonal sub-block scan for HE residual coding*, JCTVC-G323, 7th Joint Collaborative Team on Video Coding (JCT-VC) Meeting, Geneva, Switzerland, Nov. 2011.

[13] *HEVC Test Model, HM 16.6*. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6/.

[14] F. Bossen, *Common test conditions and software reference configurations*, JCTVC-L1100, 12th Joint Collaborative Team on Video Coding (JCT-VC) Meeting, Geneva, Switzerland, Jan. 2013.