Xianfeng Ou, Long Yang, Guoyun Zhang, Longyuan Guo, Jianhui Wu, Bing Tu

Improved Adaptive Transform for Residue in H.264/AVC Lossless Video Coding

The H.264/AVC was designed mainly for lossy video coding, the lossless coding of H.264 use bypass mode for DCT and quantization. Although sample-by-sample DPCM improves performance of coding, the benefit is limited in intra. In this paper, a new adaptive transform is proposed based on the character of 4×4 block residual coefficient's distribution, which can be used both in intra and inter coding. The greatest strength of the proposed transform is the decorrelation without inflation versus dynamic range of input matrix. Due to the random distribution of residual coefficients, a specific transform is hard to play a positive impact on them. Therefore, several transforms of different directions will be implemented simultaneously, and the most efficient one will be determined by a proposed mechanism. Then, by means of statistic method, a new scan order is designed for CAVLC entropy encoder, cooperating with corresponding transform. The simulation results show that based on the fast algorithm of proposed method, the bit saving achieves about 7.41% bit saving in intra coding and 10.47% in inter, compared with H.264-LS.

Key words: H264/AVC, lossless video coding, adaptive transform, residual coefficient's distribution

Poboljšana adaptivna pretvorba ostataka prilikom H.264/AVC video kodiranja bez gubitka kvalitete. H.264/AVC je napravljen prvenstveno za kodiranje videa uz gubitak kvalitete, dok kodiranje H.264 bez gubitka kvalitete koristi zaobilazni mod za DCT i kvantizaciju. Iako uzorak-po-uzorak (DPCM) kvantizacija poboljšava performanse kodiranja, dobitak je ograničen. U ovom radu predlaže se nova adaptivna transformacija koja se zasniva na znakovima od 4×4 blokova distribucije ostataka koeficijenata, koja može koristiti i unutarnje i među kodiranje. Najveća snaga predložene transformacije je u nekoreliranosti bez inflacije protiv dinamičke veličine ulazne matrice. Radi slučajne distribucije ostataka koeficijenata, teško je postići da određena transformacija ima pozitivan učinak na njih. Iz tog razloga istovremeno je implementirano nekoliko transformacija različitih pristupa, te je korištenjem predloženog mehanizma odabrana najefikasnija. Zatim je, korištenjem statističke metode, dizajniran novi poredak snimanja za CAVLC entropijski enkoder, koji surađuje s odgovarajućom transformacijom. Rezultati simulacija pokazuju da korištenjem brzog algoritma predložene metode dolazi do smanjenja korištenih bitova od 7.41% kod među kodiranja i 10.47% prilikom unutarnjeg kodiranja u usporedbi s H.264-LS.

Ključne riječi: H264/AVC, video kodiranje bez gubitka kvalitete, adaptivna transformacija, distribucija ostataka koeficijenata

1 INTRODUCTION

The H.264/MPEG-4 AVC [1-4] is an excellent international video coding standard. It was jointly developed by the Video Coding Experts Group (VCEG) of the ITU-T and the Moving Picture Experts Group (MPEG) of ISO/IEC. Experiments have shown that H.264/AVC can achieve a bit rate reduction approximately 30%–70% when compared with previous video coding standards such as MPEG-4 Part 2, H.263, H.262/MPEG-2 Part 2, etc., while providing the same or better image quality [5].

Since the H.264/AVC standard was originally designed for lossy coding, the lossless video coding was paid less attention. However, the lossless video is widely used in many fields, such as medical imaging [6], automotive vision, web collaboration [7], etc. Thus, to improve the performance of lossless coding, pulse-code modulation (PCM) has been firstly added in H.264/AVC standard. Later, the so-called fidelity range extensions (FRExt) [8] were included, which improved the efficiency of lossless encoding greatly. In the version of FRExt, a bypass mode for DCT and quantization processes is used for lossless encoding, and the residual coefficients are encoded by entropy coder directly, no matter in intra or inter. To further improve the coding performance of lossless video coding,

more and more techniques are emerging and maturing.

Several studies [10-18] have improved lossless coding based on H.264/AVC. Since H.264/AVC includes two kinds of coding—intra and inter, all studies can be divided into two groups, one is for intra coding; another can be expended to inter coding. Studies in [10-14] are intra improvements for lossless coding. Lee et al. [10] presented a vertical and horizontal spatial intra prediction method to reduce redundancy by using sample wise differential PCM (DPCM), which provided a better performance in lossless coding, and it's adopted as a part of the new draft amendment for the H.264/AVC standard [9]. Using the statistics of residual pixel values that obtained from lossless intra prediction, Heo et al. [11] proposed an improved context-based adaptive variable length coding (CAVLC) scheme for lossless intra coding by modifying the relative entropy coding in H.264/AVC. Song et al. [12] proposed a pixel-wise spatial interleave prediction to perform multidirectional intra prediction. Kim et al. [13] showed an efficient two-layered residual coding scheme, which improved the compression efficiency of H.264 intra lossless coding significantly. Gu et al. [14] proposed a mode-dependent template method for intra lossless coding by eliminating residual redundancy according to corresponding template, at the same time a new scan order for each template was designed to achieve considerable performance improvement for intra lossless coding.

Aforementioned studies were concentrated on intra coding. To achieve higher compression efficiency for video sequences, some studies [15-17] were focused on temporal redundancies. Wei et al. [15] proposed an adaptive truncation algorithm for Hadamard transformed H.264/AVC lossless video coding. In their paper, Hadamard transform was first applied to residual coefficients, no matter intra or inter. Then, based on adaptive truncation, all the residual coefficients were separated into quotient and remainder parts. Kim et al. [16] redesigned entropy coding modules due to the statistics of residual pixel values both in intra and inter, making it more efficient for lossless coding. Zhang et al. [17] proposed that intra and inter prediction residuals are in forms of granular noise, which could be further predicted and entropy coded uniquely, then the Exp-Golomb encoder was used for the prediction errors of residual image. The quotients were encoded directly, and the remainders were encoded as fixed-length after redundancies were removed according to the feature of Hadamard transform [18]. While H.265/HEVC[19][20] is able to obtain 50% higher coding efficiency than H.264/AVC, that is to say, 50% bit rate saving when the same image quality acquired, but the encoding complexity is much higher[21].

In this paper, an improved adaptive transform for residue in H.264/AVC lossless coding is proposed. In the proposed method, multi-direction transforms for prediction residue are used, to further remove correlation which is inside a 4×4 block. However, only one of proposed transform is selected adaptively. The transformed residue can be separated into DC and AC parts. The DC part is achieved adaptively from the prediction residual block, and the AC part can be considered as the prediction of neighbor pixel value. After proposed transform, several new scan orders corresponding to the direction of transform are adapted to improve the efficiency of entropy coding.

The rest of this paper is organized as follows. In Section II, we briefly review the lossless coding of H.264/AVC. Section III describes the proposed adaptive transform for residue method in detail. Experimental results validating the effectiveness of proposed algorithm are shown in Section IV. Finally, conclusions are drawn in Section V.

2 LOSSLESS VIDEO CODING OF H.264/AVC

In Lossy coding of H.264/AVC, spatial and temporal redundancies are removed by means of intra and inter prediction separately. Then the predicted residues are further transformed, quantized and then encoded by entropy encoder. Whereas, in the lossless coding, transform and quantization are bypassed to avoid amplifying the Dynamic Range (DR) of transformed coefficients and producing rounding errors. Then, the residue was directly entropy encoded until Lee et al. [10] presented the DPCM for vertical and horizontal spatial prediction residue.

According to Lee's proposition, the predicted residual image, which is denoted by $R_{(i,j)}$, can be expressed as follows:

$$R_{(i,\,j)} = P_{(i,\,j)} - I_{(i,\,j)} \qquad (1)$$

Where $P_{(i,j)}$ and $I_{(i,j)}$ are predicted image and original image, respectively. For a 4×4 luma block, when the vertical prediction is used in intra prediction, the residue of sample-by-sample (SbS) DPCM $R'_{(i,j)}$ will be:

$$R'_{(i,\,j)} = R_{(i,\,j)} - R_{(i,\,j-1)} \quad (0 \le i \le 3,\ 0 \le j \le 3) \qquad (2)$$

If the horizon prediction is used in intra prediction, the residue of SbS DPCM $R'_{(i,j)}$ will be:

$$R'_{(i,\,j)} = R_{(i,\,j)} - R_{(i-1,\,j)} \quad (0 \le i \le 3,\ 0 \le j \le 3) \qquad (3)$$

The refined horizontal and vertical prediction modes improve the efficiency of intra encoding. But for other directions of intra prediction and inter prediction, it will be helpless. Because of the bypass mode for DCT, the minimum unit of Motion Estimate (ME) can only remove the temporal redundancies, failing to get rid of the redundancies inside a unit of Motion Estimate.

Wei first applied Hadamard Transform (HT) to residual coefficients in lossless H.264/AVC [15]. There are two main purposes to use transform in video coding: one is the energy concentration, which concentrates DC component into top-left corner; another is its capability of decorrelation. However, the determinant of Hadamard matrix is 16, which means 16 times of DR for residual coefficients will be inflated after HT. If transformed coefficients are entropy encoded directly, the final code stream will also be inflated. That is why transformation is bypassed in FRExt. So, Wei proposed an adaptive truncation for HT, resulting in restrained inflation. But the truncated parts are encoded as Fixed Length (FL) coding. Sometimes, FL is less effective; even redundancy exists, it still cannot be eliminated. Take a 4×4 block HT for example:

$$Y = HXH^T \tag{4}$$

where H is a 4×4 Hadamard matrix, which is denoted in (5). X represents an input 4×4 block, which is defined particularly in (6).

$$H = H^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} (5) \tag{5}$$

$$X = \begin{bmatrix} 128 & 128 & 128 & 128 \\ 128 & 128 & 128 & 128 \\ 128 & 128 & 128 & 128 \\ 128 & 128 & 128 & 128 \end{bmatrix} \tag{6}$$

Hence:

$$Y = \begin{bmatrix} 2048 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{7}$$

According to Wei's proposition, Y can be decomposed as truncated coefficients T and FL coding F, which can be expressed as:

$$T = \begin{bmatrix} 128 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

T can be further compressed by entropy encoder. However, F is encoded directly without being compressed. Thus, if the transform is performed without inflation of DR, the truncation is unnecessary. In the next section, we will give our proposed algorithm.

3 PROPOSED ADAPTIVE TRANSFORM FOR RESIDUE

In this section, an adaptive transform for residue in H.264/AVC lossless coding is given in detail. As we discussed in the last section, if the transform is implemented without inflation, and the decorrelation will also be performed simultaneously, it would make a great deal of difference. Therefore, Hadamard matrix is substituted by the proposed matrix S_1, which denote in (10), and the transform is shown in (11).

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{10}$$

$$Y_1 = S_1 X S_1^T = \begin{bmatrix} 128 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

The biggest advantage of S_1 is no inflation of DR after transform, compared to the Hadamard transform. Therefore, there is no need of truncation and FL coding. However, this kind of matrix is not unique, and could not be available for all input residue matrix. But, if the character of prediction residues is known beforehand, such as texture, then the specific transform will be developed. In the following section, we will give a detail on how it works in lossless H.264/AVC.

The procedure of our proposed algorithm is shown in Fig.1, and the algorithm can be separated into three parts— residue analyzer, adaptive transform for residue and scan mode decision. The purpose of first two parts is to remove special correlation; of the last part is to improve the efficiency of CAVLC entropy encoder.
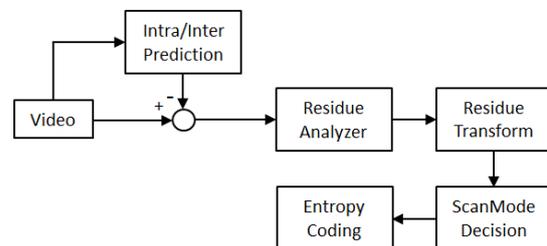

Fig. 1: Procedure of the proposed algorithm.

3.1 Residual Analyzer

According to (10) and (11), the proposed matrix S_1 produced considerable results. However, the following

S_2, S_3 and S_4 which defined in (12), (14) and (16) will achieve the same purpose (13), (15), (17).

$$S_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (12)$$

$$Y_2 = S_2 X S_2^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 128 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (13)$$

$$S_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (14)$$

$$Y_3 = S_3 X S_3^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 128 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (15)$$

$$S_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (16)$$

$$Y_4 = S_4 X S_4^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 128 \end{bmatrix} \quad (17)$$

In fact, the "128" in (8) represents the mean value of matrix X, which is different from "128" in (11), (13), (15) and (17). The latter are elements of matrix X, which are decided by S_1, S_2, S_3, and S_4. If the "128" in formula (8) is called DC component, the "128" in formula (11), (15), (16) and (17) can be called Pseudo-DC component (PDC). At the end of entropy encoding stage, because of PDC is encoded directly, the value of PDC is expected to be as small as possible. Therefore, an efficient mechanism is proposed to decide the right PDC.

The formulas indicated (11), (15), (16) and (17) are 2-D transform, which can be decompose into row transform and column transform. In row transform, S_1, S_2, S_3 and S_4 decide which column is PDC; whereas in column transform, S_1, S_2, S_3 and S_4 decide which row is PDC. According to (1), we denote RS as:

$$RS_i = \left|R_{(i, \ 0)}\right| + \left|R_{(i,1)}\right| + \left|R_{(i,2)}\right| + \left|R_{(i,3)}\right| \quad (0 \le i \le 3) \quad (18)$$

$$RS_j = \left|R_{(0,j)}\right| + \left|R_{(1,j)}\right| + \left|R_{(2,j)}\right| + \left|R_{(3,j)}\right| \quad (0 \le j \le 3) \quad (19)$$

hence

$$i = min\{RS_i\} \quad (0 \le i \le 3) \quad (20)$$

$$j = min\{RS_j\} \quad (0 \le j \le 3) \quad (21)$$

Formula (18) and (19) calculate RS_i and RS_j with different i and j. Formula (20) and (21) are used to calculate the value of i and j, namely, the corresponding value of i and j when RS_i and RS_j achieve to the minimum. Then i-row or j-column is taken as PDC-row or PDC-column. The right PDC is express as:

$$PDC = X_{(i,j)}, \quad (0 \le i \le 3, \ 0 \le j \le 3) \quad (22)$$

Fig. 2: Chose the right PDC.

Fig. 2 shows the right PDC is chosen in a 4×4 block. From Fig. 2, we infer that S_1 and S_2 are used as row transform and column transform separately. Meanwhile, the first column and fourth row are taken as column-PDC and row-PDC, and the right PDC is in the cross position.

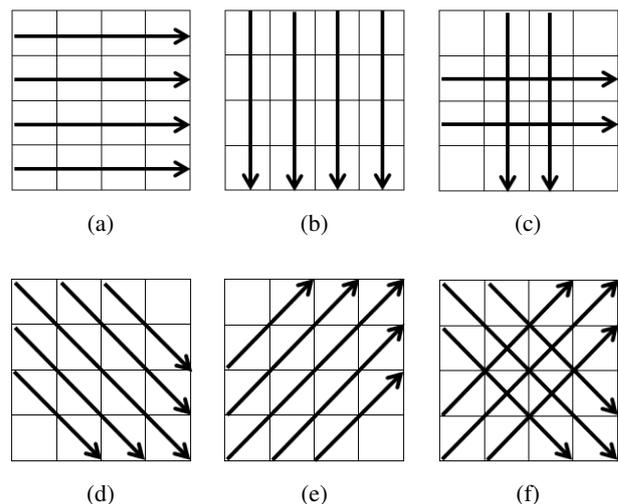### 3.2 Adaptive Transform for residue

Fig. 3: 6-Direction transforms: (a) Horizontal transform (b) Vertical transform (c) 2-D transform (d) Horizontal down transform (e) Horizontal up transform (f) Cross transform
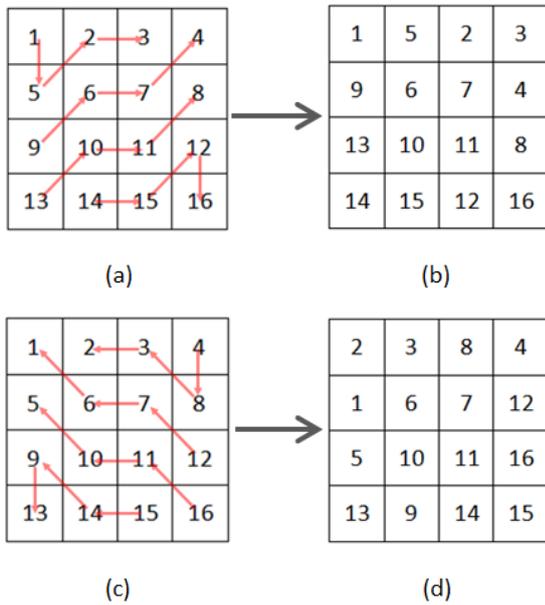
Fig. 4: Proposed reconstruction: (a), (c) Original matrix (b) Horizontal up reconstruction (d) Horizontal down reconstruction
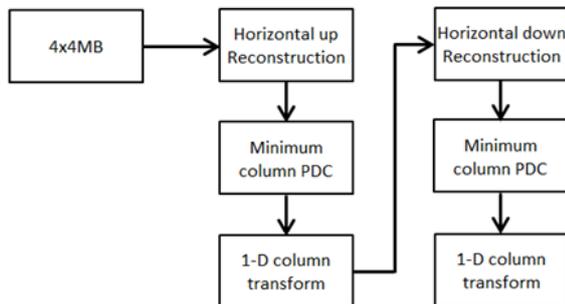

Fig. 5: Procedure of cross transform

There are 6 direction transforms are defined in our proposed algorithm, which are showed in Fig. 3.

In Fig. 3, (a), (b) and (c) are the regular 1-D and 2-D transform, which are aiming at remove correlation of horizontal and vertical directions, or both of two directions. (d) and (e) are also 1-D transform, which are used to remove correlation of horizontal down and horizontal up directions. (f) is 2-D transform, which is combined by (d) and (e). To achieve transform of (d) and (e), two reconstructed matrixes are proposed based on the new order of original matrix. Fig.4 shows the principle of proposed reconstruction.

In Fig. 4, the elements of original matrix (a) and (c) are rearranged according to the new order, with results showed in (b) and (d) separately. In (a) and (c), every four carefully selected elements are stringed together by the red arrow, composing a new row of reconstructed matrix. The stringed four elements are given full consideration of the direction of correlation. For example, the horizontal up direction is paid more attention in (a), whereas the horizontal down direction is paid more attention in (c). Thus, the horizontal direction of reconstructed matrix (b) and (d) is provided with correlation of horizontal up direction and horizontal down direction of the (a) and (c). Therefore, a column transform is performed on (b) and (d) by S_1, S_2, S_3 or S_4, which depend on (20), to achieve the horizontal up and horizontal down transform.

The cross transform in Fig. 3 (f) is combined by horizontal up transform and horizontal down transform, which means twice reconstruction are needed. Fig.5 shows the procedure of cross transform.

The above paper mainly elaborates the proposed 6 kinds of transform, which are used for removing 6 directions correlation.

However, according to (1), residues are the results of the prediction for original image, and there is a low correlation inside a residue 4×4 block. Therefore, it's hard to tell which direction needs to be transformed. Sometimes, none of them need transforming since there may be no correlation inside a residue 4×4 block. To solve these problems, a novel mechanism is proposed to choose the most efficient transform.

The six kinds of transform can be numbered from 1 to 6, according to the order of (a) to (f) in Fig.3. In addition, number 0 represents the non-transformed residue. Hence, according to formula (1), the non-transformed residue of a 4×4 block is denoted by R^0, and the 6 kinds of transformed residue are represented by R^1, R^2, R^3, R^4, R^5 and R^6. Then the sum of absolute value of a 4×4 block residue can be expressed as follows:

$$SUM\ R^m = \sum_{i=0}^{3} \sum_{j=0}^{3} \left| R_{(i,\ j)}^m \right| \ , \quad (0 \le m \le 6) \qquad (23)$$

Where m is defined as:

$$m = min\{SUM\ R^m\} \ , \qquad (0 \le m \le 6) \qquad (24)$$

Like (20) and (21), formula (25) shows that the value of m is determined by the minimum $SUM\ R^m$. Therefore, the value of m is used to decide which kind of transform will be performed on the prediction residue.

3.3 Scan Mode Decision

After transform, entropy coding is implemented. For CAVLC coder, some new scan order is proposed to reorder zero coefficients to the end of the scanned sequence.
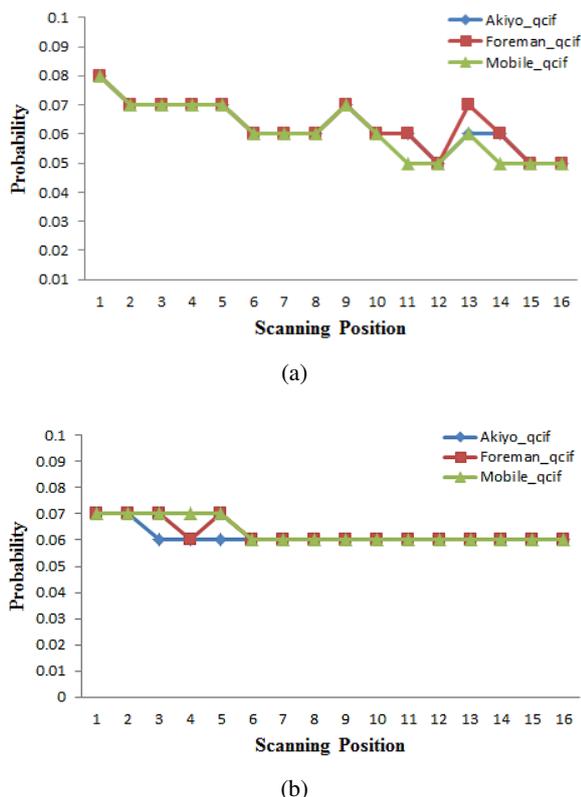
(a)

(b)

Fig. 6: Probability distribution of non-transformed zero coefficients versus scanning position

16	15	14	13
12	11	10	9
8	7	6	5
4	3	2	1

(a)

16	12	8	4
15	11	7	3
14	10	6	2
13	9	5	1

(b)

16	12	11	10
15	9	8	7
14	6	5	4
13	3	2	1

(c)

16	12	7	11
15	6	10	5
14	4	9	3
13	8	2	1

(d)

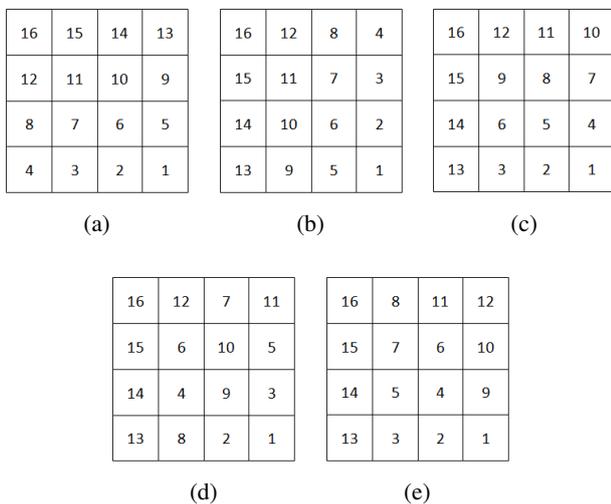| 16	8	11	12
15	7	6	10
14	5	4	9
13	3	2	1

(e)

Fig. 7: Scan position for each kind of transform (a) Non-transformed or Horizontal transform; (b) Vertical transform; (c) 2-D Transform; (d) Horizontal down Transform or Horizontal up Transform; (e) Cross Transform

Therefore, the probability distribution of zero coefficients versus the scanning position will be more important for de-

signing a new order. At first, the non-transformed residue 4×4 blocks is countered both in all-I and IPPP frames of H.264-LS video coding. Fig.6 shows the results of three representative video sequences.

As observed, the probability distribution of zero coefficients in all-I frame of H.264-LS coding is relatively uniform and tends to decrease, which conform to the characteristic of intra prediction—position "1" has high correlation of adjacent pixel, compared to low correlation of position "16". Whereas, in IPPP frame of H.264-LS coding, the probability distribution is more independent of the scanning position than all-I frame, and is hard to reorder zero coefficients to the end of scanned sequence. However, according to the proposed 6 kinds of transform, the distribution of transformed residues has turn out some regulation, which is shown as fig.10 in section 4.

Based on the statistics of probability distribution of zero coefficients, we follow the principle of scanning position prior of the low probability distribution, and later of the high probability distribution. Therefore, according to different of probability distribution, a new scan sequence is proposed for each kinds of transform block, including non-transformed block. Fig.7 shows the proposed scan sequence.

4 EXPERIMENTAL RESULTS

In this section, five 4:4:4 QCIF and three 4:4:4 1080p video sequences are tested based on JM 15.1 for showing the effectiveness of proposed method. In our experiments, the entropy coding is CAVLC to compare our compression performance with H.264-LS (CAVLC) of Lee's [10], and Wei's method [15].

The CAVCL part is implemented without change, however, we have to add some syntax elements in bitstream for decoding. For each 4×4 block, after intra/inter prediction, 3-bits syntax elements are added to indicate which proposed transform is used, since there are six kinds of transform and one kind of non-transform block. Then followed by the code of transform matrix, row transform is ahead, and column transform is later, and there is no need for non-transformed of 4×4 blocks. The 2-bits code is enough to indicate 4 different kinds of transform matrix, consisting S_1 to S_4. Table.1 shows the format of the added syntax elements. The syntax elements can be coded as fixed length, Fig.8 shows the Bitstream of syntax elements.

Table 1: Format of Syntax Elements

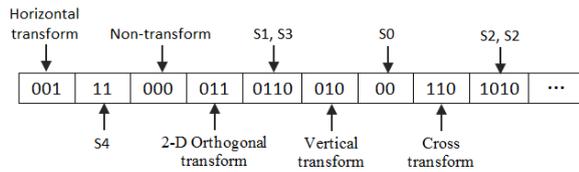| Transform type	000	001	010	011	100	101	110
suffix	none	2bits	2bits	4bits	2bits	2bits	4bits

Fig. 8: Bitstream of syntax elements

In fact, the more kinds of transform will be used, the higher of compression will be achieved since the transform of different directions will lead to further excavate potential correlation. However, more kinds of transform will also cause coding complicated and time consuming. Therefore, in Table II and III, the coding results of proposed method are given both in 3 kinds of transform and 6 kinds of transform, which named as Algorithm 1 and Algorithm 2. The 3 kinds of transform are horizontal, vertical and 2-D transform. Because the increased coding time of algorithm 1 is accepted, this can be seen from Fig.9.

The QCIF sequences of *Akiyo, coastguard, foreman, mobile and news* are coded both in 30 frame of all-I and IPPP structure. Fig.10 shows the results of probability distribution for each kind of transformed residues block. The results of transform indicates that all-I and IPPP frames have the same probability distribution for the same transform. In horizontal transform, the highest probability appears at position "1", "2", "3" and "4", and others are the same. In vertical transform, the highest probability appears at position "1", "5", "9" and "13", and the others are the same. In 2-D transform, the highest probability appears at position "1", followed by "5", "9" and "13", and then is "2", "3" and "4", and the others are the lowest. Transform of horizontal down and horizontal up have the same probability distribution, whose highest probability appears at position "1", "5", "9" and "13", then followed by "2", "4", "7", "11", "14" and "16", and the rest are the lowest. Cross Transform is similar with horizontal down or horizontal up. and all the experiments were conducted based on the Intel(R) Core(TM) i3 M330 @ 2.13GHz, 8G DDR3-1333 SDRAM and window7 64-bit operation system.

Table 2 shows that intra compression performance of proposed 3 kinds of transform (Algorithm 1) achieves about 7.41% bit saving, compared with Wei's 6.24% on average. The bit rate saving in the table is calculated by the following formula.

$$\text{Bit Saving} (\%) = \frac{\text{Bitrate}_{\text{H.264-LS}} - \text{Bitrate}_{\text{Proposed}}}{\text{Bitrate}_{\text{H.264-LS}}} \times 100 \tag{25}$$

Obviously, the improved compression performance of the proposed 3 kinds of transforms for intra coding is not satisfactory, especially in high resolution of 1080p se-
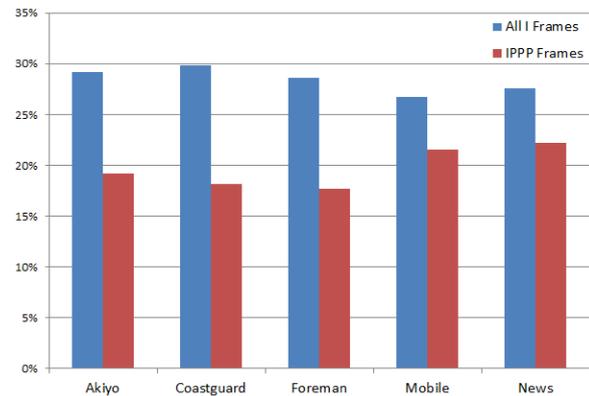

Fig. 9: Coding time increase compared with H.264-LS in percentage

quences. However, the proposed 6 kinds of transforms (Algorithm 2) give even higher performance of 8.85%, which imply that the last 3 kinds of transform, horizontal down transform, horizontal up transform and cross transform play an important role in intra coding. It's worth mentioning that the increasing time of proposed algorithm 1 will be no more than 30% compared with H.264-LS.

5   CONCLUSION

This paper proposes an adaptive residue transform algorithm for H.264/AVC lossless video coding. The proposed algorithm performed on 4×4 block predicted residue, and transformed by proposed 4 kinds of matrix. According to the character of residual coefficients, the 4 kinds of matrix can be chosen adaptively to implement several transforms with different directions of predicted residual coefficients. Finally, only one transform is determined by another proposed efficient mechanism, which is used to find out the most redundant among the proposed 6 kinds of directions. After proposed transform, the statistics feature of 4×4 blocks residual coefficients present some regulations. Based on the regulation of different transforms, the corresponding scan orders are redesigned so that the zero coefficients are arranged to the end of the scanned sequence. The simulations show that the redesigned scan orders work well with the proposed transform. With no more than 30% time increments in intra coding and 25% in inter. The proposed algorithm achieves about 7.41% bit saving in intra coding, and 10.47% in inter compared with H.264-LS. In accordance with proposed method, the bit saving can be further improved without considering the time increment, approximately 8.85% in intra and 12.60% in inter.

Table 2: Comparison of compression performance for existing and proposed algorithms (all-I frames sequences)

All-I Frames Sequence	H.264-LS [9]	Wei [15]		Proposed 3 transform (Algorithm 1)		Proposed 6 transform (Algorithm 2)	
	Rate (Mb)	Rate (Mb)	Saving (%)*	Rate (Mb)	Saving (%)*	Rate (Mb)	Saving (%)*
Akiyo**	6.86	6.32	7.81	**6.13**	**10.64**	6.01	12.39
Coastguard**	8.14	7.68	5.74	**7.55**	**7.25**	7.44	8.60
Foreman**	7.81	7.17	8.23	**7.01**	**10.24**	6.86	12.16
Mobile**	11.95	10.97	8.16	**10.85**	**9.21**	10.73	10.21
News**	7.74	7.16	7.50	**6.96**	**10.01**	6.86	11.37
ducks***	308.05	294.64	4.35	**294.43**	**4.42**	290.46	5.71
In To Tree***	273.75	**259.92**	**5.05**	262.35	4.16	258.44	5.59
Town Cross***	293.78	**284.01**	**3.33**	284.03	3.31	279.88	4.73
Average			6.24		**7.41**		8.85

*Bitsaving calculated relative to H.264-LS.

**30 frames simulated with a frame size of 176.

***10 frames simulated with a frame size of 1920×1080.

Table 3: Comparison of compression performance for existing and proposed algorithms (IPPP frames sequences)

IPPP Sequence	H.264-LS [9]	Wei [15]		Proposed 3 transform (Algorithm 1)		Proposed 6 transform (Algorithm 2)	
	Rate (Mb)	Rate (Mb)	Saving (%)*	Rate (Mb)	Saving (%)*	Rate (Mb)	Saving (%)*
Akiyo**	2.95	2.94	0.17	**2.74**	**7.12**	2.66	9.83
Coastguard**	7.18	6.93	3.41	**6.78**	**5.57**	6.65	7.38
Foreman**	7.13	7.03	1.46	**6.53**	**8.42**	6.37	10.66
Mobile**	10.25	9.65	5.89	**9.41**	**8.20**	9.28	9.46
News**	4.16	4.14	0.30	**3.79**	**8.89**	3.69	11.30
ducks***	341.44	324.60	4.93	**297.73**	**12.80**	290.80	14.83
In To Tree***	311.75	293.01	6.01	**262.22**	**15.89**	255.30	18.11
Town Cross***	345.29	319.68	7.42	**287.07**	**16.86**	278.84	19.24
Average			3.70		**10.47**		12.60

*Bitsaving calculated relative to H.264-LS.

**30 frames simulated with a frame size of 176×144.

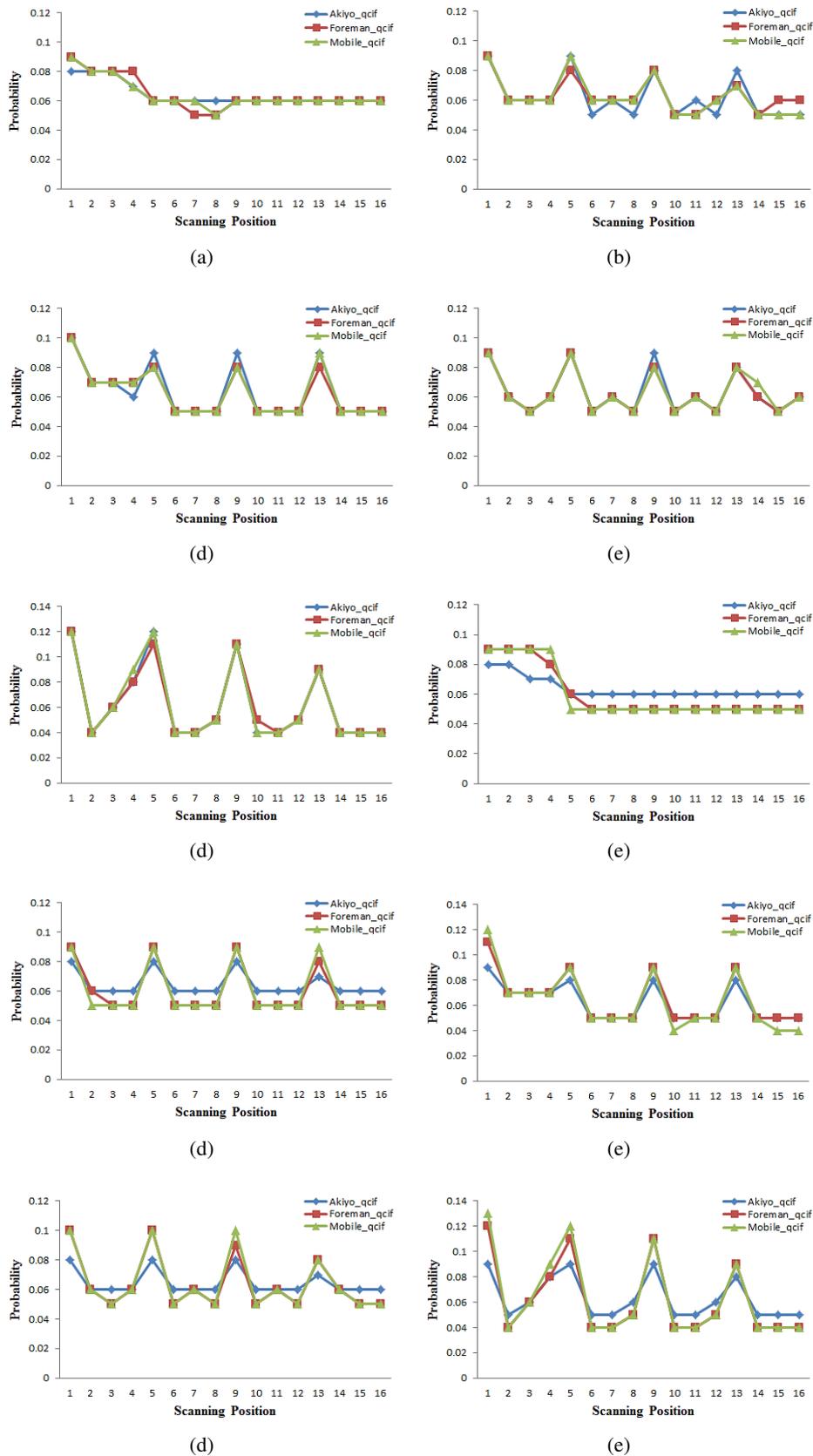***10 frames simulated with a frame size of 1920×1080.

6 ACKNOWLEDGEMENTS

Fig. 10: Probability distribution of zero coefficients versus scanning position (a)(b)(c)(d)(e) All-I frame, qcif; (f)(g)(h)(i)(j) IPPP frame, qcif;(a) (f) Horizontal Transform; (b)(g) Vertical Transform; (c)(h) 2-D Transform; (d)(i) Horizontal down Transform or Horizontal up Transform;(e)(j) Cross Transform

REFERENCES

[1] Advanced Video Coding for Generic Audiovisual Services, ITU-T Rec.H.264 and ISO/IEC 14496-10 AVC, version 3, 2005.

[2] A. Luthra, G. J. Sullivan, and T. Wiegand, "Introduction to the special issue on the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 557–559,Jul. 2003.

[3] T. Wiegand, G. J. Sullivan, G. Bjntegaard, 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.

[4] G. J. Sullivan and T. Wiegand, "Video compression—from concepts to the H.264/AVC standard," *Proc. IEEE*, no. 1, pp. 18–31, Jan.2005.

[5] *Lossless Coding and QP Range Selection, document* JVT-C023, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, May 2002.

[6] D.Clunie,"Lossless compression of grayscale medical images-effectiveness of traditional and state of the art approaches," *Proc. SPIE Med. Imag.*, vol. 3980, pp. 74–84, Feb. 2000.

[7] M. H. Zhou, W. Gao, M. Q. Jiang, and H.P. Yu, "HEVC Lossless Coding and Improvements," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1839-1843, Dec 2012.

[8] G. J. Sullivan, T. McMahon, T. Wiegand, and A. Luthra, Eds., *Draft Text of H.264/AVC Fidelity Range Extensions Amendment to ITU-T Rec*. H.264 j ISO/IEC 14496-10 AVC, ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16 Joint Video Team document JVT-L047, Jul. 2004.

[9] *Draft Text of H.264/AVC Advanced 4:4:4 Profile Amendment to ITU-TRec*. H.264|ISO/IEC 14496-10 AVC, Standard ISO/IEC JTC1/SC29/WG11 and ITU-T Q.6/SG16, Oct. 2005.

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

[11] J. Heo and Y. S. Ho, "Efficient level and zero coding methods for H.264/AVC lossless intra coding," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 87–90, Jan. 2010.

[12] L. Song, Z. Y. Luo, and C. Xiong "Improving Lossless Intra Coding of H.264/AVC by Pixel-wise Spatial Interleave Prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 12, pp. 1924–1928, DEC. 2011.

[13] S. H. Kim, J. W. Kang, and C.-C. J. Kuo, "Improved H.264/AVC lossless intra coding with two-layered residual coding (TRC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 7, pp. 1005–1010, Jul. 2011.

[14] Z. Y. Gu, W. S. Lin, B. S. Lee, C. T. Lau and M. T. Sun, "Mode-Dependent Templates and Scan Order for H.264/AVC-Based Intra Lossless Coding," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4106–4116, Sep. 2012.

[15] S. T. Wei, C. W. Tien, B. D. Liu, and J. F. Yang, "Adaptive Truncation Algorithm for Hadamard-Transformed H.264/AVC Lossless Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 538–549, May 2011.

[16] S.-H. Kim, J. Heo, and Y.-S. Ho, "Efficient entropy coding scheme for H.264 lossless video coding," *Signal Process.*, Image Commun., vol. 25, no. 9, pp. 687–696, Oct. 2010.

[17] Q. Zhang, Y. Dai, and C. C. J. Kuo, "Lossless video compression with residual image prediction and coding (RIPC)," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2009, pp. 617–620.

[18] T. Mochizuki, "Bit pattern redundancy removal for Hadamard transformation coefficients and its application to lossless image coding," Electronics and Communications in Japan, Part3, Vol. 80, No.6, 1997.

[19] A. Alshin. Coding efficiency improvements beyond HEVC with known tools[C]. SPIE Optical Engineering + Applications. International Society for Optics and Photonics, 2015.

[20] D. Flynn, D. Marpe, M. Naccari, et al. Overview of the Range Extensions for the HEVC Standard: Tools, Profiles and Performance[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015:1-1.

[21] D. Marquardt, M. Jongbloed-Pereboom, B. Staal, et al. Performance comparison of H.265/MPEG-HEVC, VP9 and H.264/MPEG-AVC encoders[C]. 2013 Picture Coding Symposium (PCS). 2013:e954.

X. Ou received the B.S. degree in Electronic Information Science and Technology and M.S. degree in Communication and Information System from Xinjiang University, Xinjiang, China, in 2006 and 2009, respectively. In 2016, he received his Ph.D. degree in Communication and Information System from Sichuan University, Chengdu, China. He was a visiting researcher at the Internet Media Group, Politecnico di Torino, Turin, Italy, from Jan. to Apr. 2014, working on distributed video coding and transmission. His main research interests include image and video coding technologies.

L. Yang received his B.S degrees from Shanxi University of Science and Technology, Shanxi, China, in 2004, and his M.S degree in University of Electronic Science and Technology, Chengdu, China, in 2010. In 2014, he received his Ph.D. degree in the College of Electronics and Information Engineering of the Sichuan University, Chengdu, China. He is now a senior algorithm engineering in Shenzhen Intellifusion Technologies Co., Ltd. His current research interests are image processing and machine learning.

G. Zhang was born in1971. He received the Ph.D degree in pattern recognition from Hunan university in 2006. He is a professor at Hunan Institute of Science and Technology. He was a visiting researcher at George Fox University from Jan. to Jun. 2014, his research interests include image processing, computer vision and pattern recognition, etc.

L. Guo received the B.S. degree in Industrial Electrical Automation and M.S. degree in Control Theory and Control Engineering from Shenyang Ligong University, Shenyang, China, in 1996 and 2003, respectively. In 2009, he received his Ph.D. degree in Pattern Recognition and Intelligent System from Nanjing University of Science and Technology, Nanjing, China. He was a visiting researcher at the University of California, Davis, CA, USA, from Aug. 2014 to Jan. 2015, and working on Multi-views Reconstruction. His main research interests include computer vision stereo matching algorithm and image processing technology.

J. Wu received the B.S. degree in Physics from Hunan Normal University, Changsha, China, in 2000. Then he received his M.S. degree and Ph.D. degree in Optical Information Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2009, respectively. He was a visiting researcher at the Computer Vision Group, University of Nevada, Reno, American, from Aug. 2015 to Aug. 2016, working on object detection and recognition. His main research interests include object detection and object recognition.

B. Tu eceived the B.S. degree from Guilin University of Technology, Guilin, China, in2006. and Ph.D. degrees in Beijing University of Technology, Beijing, China, in2013. He is currently working toward the Postdoctoral in electrical engineering in Hunan University, Changsha, China. In 2013, he joined the College of Information and Communication Engineering, Hunan Institute of Science and Technology. From September 2013 to now, he is lecturer with the department of computer science. He is engaged in image fusion, pattern recognition, and hyperspectral image classification, image and video coding technologies.

AUTHORS' ADDRESSES

Xianfeng Ou, PH.D
Prof. Guoyun Zhang (Corresponding author)
Associate Prof. Longyuan Guo
Prof. Jianhui Wu PH.D
Bing Tu PH.D
College of Information & Communication Engineering,
Hunan Institute of Science & Technology,
Xueyuan Road, Yueyang, Hunan, China.
email: gyzhang@hnist.edu.cn

Long Yang PH.D (Co-first author, corresponding author)
Shenzhen Intellifusion Technologies Co.,
Ltd Shangbu road NO.1003, Futian district, Shenzhen, China.
email: ylong_321@126.com