A NOVEL COEFFICIENT SCANNING SCHEME FOR DIRECTIONAL SPATIAL PREDICTION-BASED IMAGE COMPRESSION

Xiaopeng Fan[1], *Yan Lu*[1], *and Wen Gao*[2]

[1]Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China
[2]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
xpfan@jdl.ac.cn, ylu@ieee.org, wgao@ict.ac.cn

ABSTRACT

Spatial prediction is a promising technique for image coding. For example, the coming AVC/H.264 standard adopts directional spatial prediction in the Intra frame coding. Similar to the traditional DCT-based image coding schemes, it still scan the transform coefficients in a zigzag order, which is inefficient for coding the residual signals predicted from the different directions. To tackle this problem, a new scheme of scanning the transform coefficients by utilizing the spatial prediction information is proposed in this paper. The distribution of transform coefficients from each direction of spatial predictions is fully studied. According to the statistics, the adaptive scan table is derived for each type of spatial predictions, which is indicated by the prediction mode. Experimental results demonstrate that the proposed scheme can always outperform the JVT codec using zigzag scanning. Moreover, it does not introduce any extra computing costs in software implementation.

1 INTRODUCTION

Image compression plays a very important role in many applications. The fundamental problem in image compression is how to efficiently exploit the spatial redundancies in an image. Transform coding developed in 1960s is the core technology for this purpose. Recently, we have seen an impressive advance in wavelet coding [1]~[3]. Although wavelet transform is capable of providing more flexible functionalities in scalable coding, DCT is still widely used in many practical coding systems because of its computational efficiency. Particularly, the coding scheme combining DCT and spatial prediction is becoming a promising technique recently. For example, the directional spatial prediction scheme has been adopted in the coming AVC/H.264 standard for Intra frame coding [4][5]. In [6], it has been shown that the AVC/H.264 Intra coding outperforms JPEG2000 at low bit rate and is slightly outperformed at middle and high bit rates. In other words, the AVC/H.264 Intra coding is among the state-of-the-art image compression algorithms.

Besides the transform, another concern in image compression is how to organize and compress the transform coefficients. The distributions of DCT coefficients of images/videos have been extensively studied in the past two decades [7][8]. The result would be helpful, for instance, in design of entropy coding and quantizer. Statistics and mathematical analysis have indicated that the AC coefficients usually resemble Laplacian distributions, and the variances of the distributions across various coefficients become smaller as we go to the higher frequency coefficients. Therefore, the zigzag scan order is suitable for 2D DCT coefficients for the purpose of energy concentration.

However, recent researches show that the zigzag scan is not optimum for DCT-based image compression. For example, some DCT-based coding schemes with the reorganization of its transform coefficients have produced higher coding efficiency rather than the zigzag scan [9][10]. These schemes were inspired by the tree structure scan developed in the wavelet-based coding [1][2]. Accordingly, in order to improve the coding efficiency of spatial prediction-based image compression, the new scan orders should also be developed. An intuitive way is to utilize the tree structure scan in spatial prediction-based image compression. However, the tree structure may not be very efficient, because the transform following the spatial prediction is usually performed on a small block size, e.g. 4x4 block in AVC/H.264.

This paper focuses on developing the more efficient scanning scheme to substitute the zigzag scanning for spatial prediction-based image coding. AVC/H.264 Intra coding is adopted as the baseline of the proposed algorithm. Since the coefficients in AVC/H.264 Intra coding is achieved from the 4x4 transform on the residual signals predicted from the different directions, the distribution of coefficients of each type of prediction may be different as well. Therefore, the distributions of the coefficients are first studied in this paper. A novel method is then derived to build the spatial prediction-based tables to scan the transform coefficients. For each direction of spatial prediction, a unique scanning table is built. The scanning table used in the encoder and decoder can be indicated by the mode of spatial prediction, i.e. the coding mode. In this way, there is no extra bit introduced in the proposed image coding algorithm.

The rest of this paper is organized as follows. Section 2 overviews the intra frame coding scheme in AVC/H.264. Section 3 presents the analysis of the distribution of coefficients derived from the spatial predicted residual signals in AVC/H.264. Section 4 presents the design of scanning order tables. The simulation results are illustrated in Section 5. Finally, Section 6 concludes this paper.

2 OVERVIEW OF JVT INTRA CODING

A novel feature of AVC/H.264 Intra coding is that spatial correlations among blocks are exploited by utilizing the directional spatial prediction. The idea is based on the observation that adjacent blocks tend to have the similar textures. Therefore, as a first step in the encoding process for a given block, one may predict the block to be encoded from the surrounding blocks (typically the blocks located on top and to the left of the block to be encoded, since those blocks would have already been encoded). In AVC/H.264, two types of intra predictions are employed: 4x4 Intra prediction and 16x16 Intra prediction. Figure 1 illustrates the nine modes of 4x4 Intra prediction, including DC prediction and eight directional modes. As shown in Figure 1, only neighboring pixels of the current block contribute to the prediction. For 16x16 intra prediction, four modes are defined in terms of four directions.

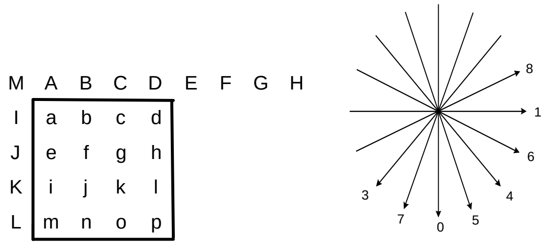


Figure 1: Intra prediction for a 4x4 block in AVC/H.264

After predicted from the neighboring blocks, the residuals signals are fed into the 4x4 integer transform, i.e.

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix}$$

The employed transform is an approximation of DCT, which is primarily 4x4 in shape and opposed to the usual floating-point 8x8 DCT specified with rounding-errors. The normalization of the transform and inverse are combined with quantization and dequantization, respectively. In AVC/H.264, the transform coefficients scanned in the zigzag order is compressed with CAVLC or CABAC entropy coder [4].

3 PROBLEM STATEMENTS

In order to achieve the distributions of the transform coefficients of the AVC/H.264 Intra frame coding, we make the statistics as follows. The mode decision depends on the mean square error (MSE) between pixels in the predicted block and the current block. The paper aims at deriving the more efficient scanning strategy for spatial prediction-based image compression, which should be quantization independent. In order to eliminate the influence of quantization in intra prediction, the original neighboring pixels are instead of the reconstructed pixels used for the spatial prediction, which can achieve the more accurate statistics without the influences of quantization. Since the quantization is not performed in the process, the transform coefficients have to be normalized by an element-by-element multiplication with a matrix, i.e.

$$\begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix},$$

where a = 1/2, and b = $\sqrt{2/5}$ as used in [11].

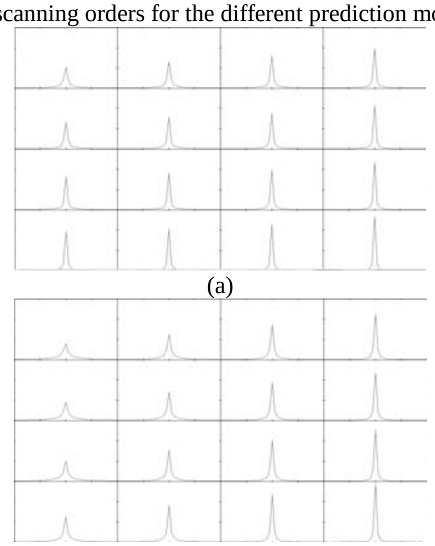
Figure 2(a) shows a group of histograms of coefficients with DC prediction mode. The position of each histogram in the figure corresponds to the position in the 2-D coefficient matrix. Figure 2(b) and Figure 2(c) shown the histograms with vertical prediction and horizontal prediction, respectively. In the figures, the top-left histograms are for the DC coefficient and the rest are for the AC coefficients. The scaling of the histogram is kept the same for all coefficients in the same figure. A taller and thinner histogram indicates the smaller variance, whereas a flatter one indicates the larger variance. Similar to the case that transform is directly applied to the raw pixels, shapes of the histograms of the higher frequency is taller and thinner than that of the lower frequency, either in horizontal or in vertical directions.

According to Figure 2, two conclusions can be drawn as follows.

1. There is no significant difference between the shapes of histograms of the DC and AC coefficients expect the width and height. Since DC coefficient is the mean of the residues, it can also be taken as a lower frequency AC coefficient.
2. In Figure 2(b), the shapes of the histograms at different positions vary steeply in horizontal direction but smoothly in vertical direction. In Figure 2(c), the situation is in the contrast.

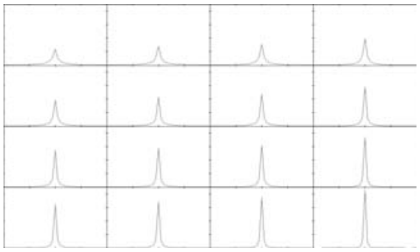
Therefore, the distributions of the transform coefficients in terms of the different prediction modes are different as well. Figure 2(a) is essentially self-symmetrical along the top-left bottom-right diagonal. Therefore, the zigzag scanning might be suitable for DC prediction. However, for Figure 2(b) and Figure 2(c), the distributions of the histograms are different in the horizontal and vertical directions. Therefore, the zigzag scanning might be inefficient for the vertical and horizontal prediction modes. The following paragraph analyzes the reasons causing the different distributions of the histograms.

In this paper, we take vertical prediction mode as an instance. In AVC/H.264, the vertical prediction mode is selected when the spatial prediction in this vertical direction is more similar to the current block rather than the other directions. The predicting block in terms of vertical prediction has the same pixel of each column. Thus, it can only reduce the horizontal correlations of the current block rather than the vertical correlation, which results in the different distribution of histograms in vertical and horizontal directions, as shown in Figure 2(b). For the other spatial prediction directions, similar properties can be found as well. Since the distribution of coefficients in terms of different prediction directions are also different, it is reasonable to use the different scanning orders for the different prediction mode.



(a)

(b)



(c)

Figure 2: Histograms of residual coefficients of different spatial prediction modes. (a) DC mode, (b) vertical mode, and (c) horizontal mode.

4. DESIGN OF SCANNING TABLES

Based on the statistic results, a new strategy for scanning the transform coefficients is presented in this section. The scan order of a position depends on the probability distribution of the coefficients at that position. Basically, the scan starts from the coefficient with the max variance, approximately ordering the coefficients in decreasing order of variance. In this way, the 2D array can be converted into a 1D vector concentrating the non-zero coefficients at the beginning and producing long runs of zeros at the end of the vector, which makes the run-length coding and the usage of the symbol of End-of-Block (EOB) very efficient.

As analyzed in Section 3, the coefficients at the same position resemble the Laplacian distribution, including the DC coefficients. A flatter Laplacian distribution indicates the larger variance and higher probability to be non-zero after quantization. This conclusion is drawn according to the property of Laplacian distributions as follows. Let:

$$f(x) = \frac{\mu}{2} e^{-\mu|x|}$$

to be the probability density, where x is the coefficient and μ is the Laplacian parameter. The uniform quantization with the offset of deadzone is used in JVT codec, as shown in Figure 3. The coefficient is quantized to be zero if it is less than the value a, which is related to the quantization step size. Thus, the probability of the coefficient to be quantized to non-zero is:

$$P(|x| \geq a) = e^{-a\mu}$$

This is a monotone descending function. A flatter Laplacian distribution has a smaller μ, i.e. it has higher probability to be non-zero after quantization, which is independent of the quantization parameter. Therefore, the table of scan order can be designed according to the Laplacian distribution. For instance, if the 16 positions in the 4x4 integer transform are sorted according to their Laplacian parameters, the scan table can be derived by comparing the variance of the distribution at each position. After the variances of all positions under each spatial prediction mode are calculated and sorted in descending, the scanning table of each mode can be derived as well. Only one unique scanning table is employed for each prediction mode, whatever the quantization parameter is used.

The following presents an example to illustrate the design of scanning table. Table 1 illustrates the variances of the distributions of 4x4 transform coefficients in terms of vertical prediction and horizontal prediction, respectively. These variances are normalized by the variance of DC coefficient at the top-left corner. Obviously, the scan orders for the two prediction modes should be different. For vertical prediction, the larger variances are distributed in the left, whereas for the horizontal prediction, the large variances are distributed in the top. Correspondingly, the derived scanning order are shown in Figure 4(a) and Figure 4(b), respectively.

In the proposed coding algorithm, these new designed tables are used instead of zigzag scanning in AVC/H.264. Since the table used in the encoder and decoder can be indicated by the prediction mode, there is no extra bit encoded into the bitstreams.

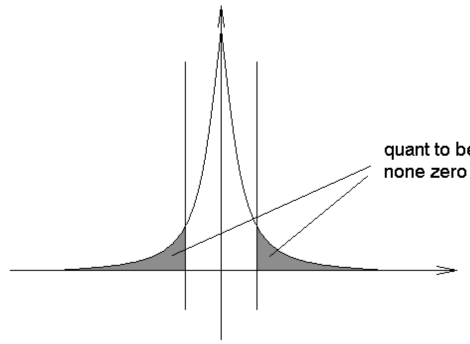
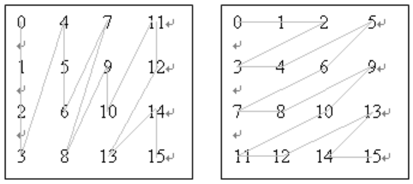


Figure 3: Quantization to the coefficients with Laplacian distribution.

Table 1: Variances of the distributions of coefficients

VERTICAL				HORIZONTAL			
1.0000	0.7573	0.5777	0.3893	1.0000	0.9254	0.8563	0.6675
0.9208	0.7101	0.5172	0.3618	0.7133	0.6871	0.5968	0.4443
0.8836	0.6387	0.4605	0.2978	0.5212	0.4707	0.4190	0.2689
0.7626	0.5406	0.3457	0.2175	0.3780	0.3473	0.2668	0.1619



(a) (b)

Figure 4: Transform coefficient scanning order of (a) vertical prediction mode, and (b) horizontal prediction mode.

5. EXPERIMENTAL RESULTS

In order to evaluate the proposed image coding scheme, some experiments have been performed. Since AVC/H.264 is adopted as a baseline in the proposed scheme, the original AVC/H.264 Intra coding is compared in this paper. The test model JM3.9 is used as the platform. Some coding conditions are set as the same. The CAVLC and CABAC entropy coding methods are tested, respectively. The image/sequences used here are illustrated in the [12], which presents the common condition of JVT testing. Since the proposed scheme is for the still image compression, each frame in the sequences is encoded as an Intra frame. In order to verify that the proposed coefficient scanning strategy is independent of the quantization parameter, the different quantization step sizes are used, i.e. 24, 28, 32 and 36.

Table 3 illustrates the experimental results using CAVLC entropy coding. The results show that the proposed scheme can averagely save about 2% bits compared to AVC/H.264 Intra coding with zigzag scanning. Using the CABAC entropy coding, the average bits saving is 1.78%. Since the AVC/H.264 Intra coding is among the state-of-the-art image compression schemes, it is very difficult to achieve any improvement. In other words, the average bit saving is significant enough. Another significant thing is that the proposed algorithm can always outperform the original AVC/H.264 Intra coding with zigzag scanning in our experiments, and meanwhile it does not introduce any computing cost in software implementation.

6. CONCLUSIONS AND FUTURE WORKS

A new transform coefficient scanning strategy for the directional spatial prediction-based image compression has been presented in this paper. Experimental results demonstrate that the proposed scheme can always outperform AVC/H.264 Intra frame coding without introducing any extra computing costs. And also, it has been shown that the zigzag scanning in AVC/H.264 is far from optimal. Since the entropy coding used in the proposed algorithm is developed for the zigzag scanning in AVC/H.264, the future work of this paper is to adjust the entropy coder according to the proposed adaptive scanning strategy. Thus, the coding efficiency can be expected to be further improved.

7. ACKNOWLEDGEMENTS

The work in part has been supported by National Hi-Tech Research Program (863) of China (2002AA119010), National Fundamental Research and Development Program (973) of China (2001CCA03300).

REFERENCE

[1] J. Shapiro, "Embedded image coding using zerotree of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3445-3463, 1993.

[2] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no 3, pp. 243-250, 1996.

[3] ISO/IEC JTC1/SC29/WG1, "JPEG2000 verification model 7.0," ISO/IEC JTC1/SC29/WG1 N1684, April 2000.

[4] T. Wiegand, "Joint final committee draft (JFCD) of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)," Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG 4th Meeting, JVT-D157, July, 2002.

[5] T. Wiegand and G. Sullivan, "The Emerging JVT/H.26L video coding standard", Tutorial at ICIP 2002, Rochester, NY, Sept. 2002.

[6] T. Hallbach, "Performance comparison: H.26L intra coding vs. JPEG2000," Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T 4th Meeting, JVT-D039, July, 2002.

[7] R. Reininger and J. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM-31, pp. 835–839, June 1983.

[8] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Processing*, vol. 9, no. 10, pp.1661-1666, Oct. 2000.

[9] Z. Xiong, O. Guleryuz, and M. Orchard, "A DCT-based embedded image coder," IEEE Signal Processing Letters, vol. 3, no. 11, pp. 289-290, 1996.

[10] D. Zhao, Y. K. Chan and Y. Lu, "Embedded image coding based on novel organization of DCT coefficients," *SPIE International Symposium on Optical Science and Technology*, vol. 4115, pp. 153-162, San Diego, Sept. 2000.

[11] A. Hallapuro, M. Karczewicz, and H. Malvar, "Low complexity transform and quantization – part I: basic implementation," Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG 2nd Meeting, JVT-B038, Feb. 2002.

[12] G. Sullivan, "Recommended simulation common conditions for H.26L coding efficiency experiments on low-resolution progressive-scan source material," ITU-T VCEG SG16 Q.6 14th Meeting, VCEG-N81, Santa Barbara, USA, Sept. 2001.

Table 2: Experimental results with CAVLC entropy coder

sequence	frames	QP	psnrY	bits/frame		bit saving
				JM3.9	Proposed	
mobile	300	24	38.44	389448	386878	0.66%
Paris	300	24	39.45	231042	227400	1.58%
tempete	260	24	39.05	249020	246120	1.16%
container	300	24	39.72	42906	42208	1.63%
news	300	24	40.42	44688	44083	1.35%
silent	300	24	39.2	45614	44804	1.78%
foreman	300	24	39.38	41108	40502	1.47%
mobile	300	28	35.03	279656	276511	1.12%
paris	300	28	36.39	161735	158567	1.96%
tempete	260	28	35.82	174266	171139	1.79%
container	300	28	36.94	29056	28487	1.96%
news	300	28	37.45	31259	30584	2.16%
silent	300	28	36.24	29737	29016	2.42%
foreman	300	28	36.65	27257	26689	2.08%
mobile	300	32	31.49	195262	191649	1.85%
paris	300	32	33.22	110982	108108	2.59%
tempete	260	32	32.51	117885	114921	2.51%
container	300	32	34.12	19599	19174	2.17%
news	300	32	34.39	21424	20885	2.52%
silent	300	32	33.33	18910	18420	2.59%
foreman	300	32	33.83	17746	17250	2.79%
mobile	300	36	28.2	129106	126139	2.30%
paris	300	36	30.27	73217	71631	2.17%
tempete	260	36	29.49	73092	71410	2.30%
container	300	36	31.3	12786	12464	2.52%
news	300	36	31.34	14226	13919	2.16%
silent	300	36	30.85	11353	11220	1.17%
foreman	300	36	31.22	10873	10602	2.49%
Average						1.97%