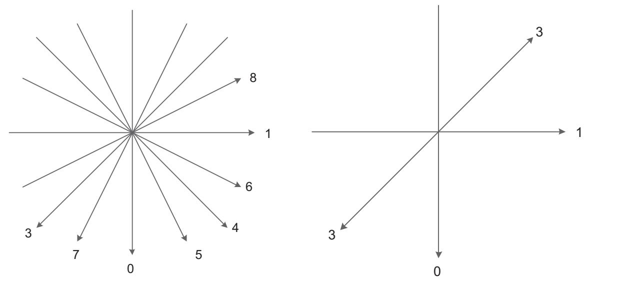
Directional Samples Reordering for Intra Residual Transform

Qingbo Wu[1], Hongliang Li[2], Tiantang Chen[3]

School of Electronic Engineering, University of Electronic Science and Technology of China
Chengdu, 611731, China

[1] wuqingbo911@gmail.com
[2] hlli@uestc.edu.cn
[3] chentiantang@163.com

***Abstract*—In this paper, a directional samples reordering (DSR) based algorithm was proposed for intra residual data transform. To make the residual more suitable for discrete cosine transform, the diagonal edge in arbitrary size of intra prediction block can be rotated to a regular horizontal edge by reordering the samples in the block. The intra residual data will be used to implement 2D discrete cosine transform after DSR procedure. Experimental results show that up to 0.5967 dB and on average 0.4372 dB gain can be achieved for CIF sequence in high bitrate with the proposed algorithm at high complexity mode than H.264 intra coding.**



(a) (b)

Fig. 1. The directional prediction modes of different intra prediction block types: (a) shows 8 directional prediction modes of 4x4/8x8 block, (b) shows 3 directional prediction modes of 16x16 block

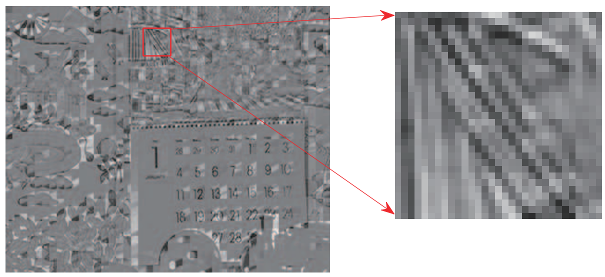


Fig. 2. Directional information in intra residual

I. Introduction

H.264/AVC [1], [2] is the state of the art international video coding standard in recent years. It has been proved that H.264 can offer much better coding efficiency than preceding video coding standard such as MPEG 1/2/4 and H.261/H.263. To achieve this performance, many new coding tools are employed in H.264 including: intra prediction based on neighboring blocks, DCT-like integer transform, in-loop deblocking filters, etc.

In the H.264, directional intra prediction technique plays an important role in reducing spatial redundancy of spatially adjacent blocks. For an intra coding macroblock, there are three intra prediction block types including: 4x4, 8x8 and 16x16 in the high profile for luma component. And different intra prediction block types have different number of prediction modes corresponding to them. For 4x4 and 8x8 intra prediction block, there are eight directional prediction modes as shown in Fig.1 (a). For 16x16 block, only three directional prediction modes are alternative as shown in Fig.1 (b).

Although the exiting directional intra prediction can efficiently reduce the bits by just coding intra residual, the potential of improving intra coding efficiency by exploiting direction information in intra residual still attracts our attention in recent studies. Since all the intra prediction values are produced from the extrapolation of neighboring reconstructed pixels [1], [3], further, the intra residual is generated from the difference of original pixels value and the synthetic prediction value. Accordingly, the intra residual will still contain abundant texture information which may be full of diagonal edges in each intra prediction block. In Fig.2, the local scene magnified from the rect area marked in intra residual is shown for us, and some significant diagonal edges can be found clearly.

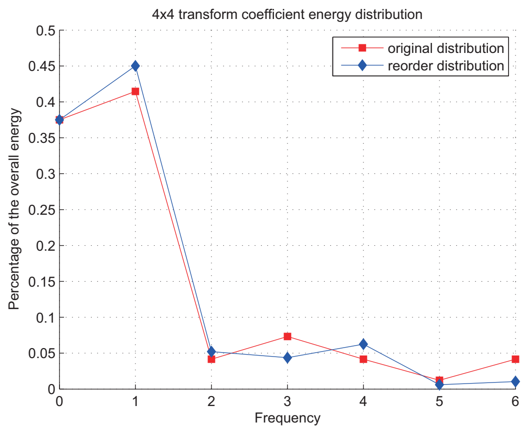
From [4], [5], we know that 2D DCT is more suitable for the condition that there are only horizontal or vertical edges in a block. Though some algorithms [7], [8], [9] have been developed to exploit directional information in original pixel dominant, while the directional information in intra residual has not been taken full advantage of in transform. In [10], a mixed spatial-DCT-based coding scheme is proposed, which focuses on improving DCT performance of inter prediction error. It separates inter prediction error into two parts through a threshold which in fact tries to guarantee the error as flat as possible in each part to improve its DCT performance. Recently, an interesting work of DDCT [5] has been proposed to improve the transform process for image coding. Even though

978-1-4577-1434-4/11/$26.00 ©2011 IEEE

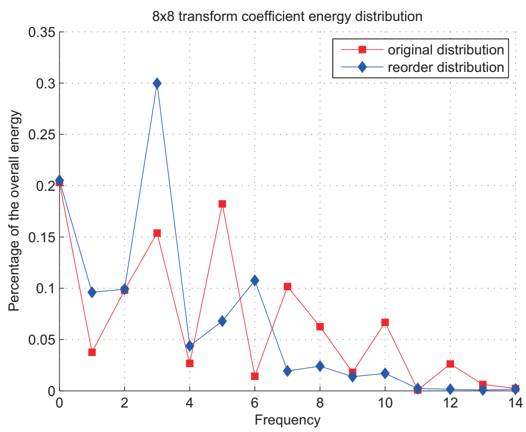
DDCT [5] can dispose similar problem by adjusting the length and sequence of each 1D DCT, it can't satisfy compatibility with H.264 framework and many DCTs used in DDCT do not have a length 2^M to facility a fast implementation.

The main contribution of our proposed method is to provide an efficient way to improve coding efficiency of intra residual full of diagonal edge with directional samples reordering (DSR). Residual samples in any size of intra prediction block (include 4x4, 8x8 or larger) are reordered according to a sort of scan mode corresponding to the edge direction in each block. As most diagonal edges in intra prediction block can be rotated to regular horizon, consequently, the proposed method achieves better performance than H.264 Intra coding at high complexity.

The rest of this paper is organized as following. In section II, we analyze the direction information of intra residual by inspecting and the transform processes in H.264. Then an efficient directional samples reordering scheme for intra residual is proposed in section III. Experimental results are shown in section IV. Finally, we conclude the paper and state our future work in section V.



(a) 4x4 transform coefficient energy distribution



(b) 8x8 transform coefficient energy distribution

Fig. 3. Energy distribution of the transform coefficients before and after reordering the samples

$$E_1 = \begin{pmatrix} 0 & 255 & 255 & 255 \\ 0 & 0 & 255 & 255 \\ 0 & 0 & 0 & 255 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{after2}-\text{DDCT}}{\Longrightarrow}$$
$$C_1 = \begin{pmatrix} 382.5 & -284.4 & 0 & -20.2 \\ 284.4 & -127.5 & -117.8 & 0 \\ 0 & 117.8 & -127.5 & -48.8 \\ 20.2 & 0 & 48.8 & -127.5 \end{pmatrix}$$
$$E_2 = \begin{pmatrix} 255 & 255 & 255 & 255 \\ 255 & 255 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{after2}-\text{DDCT}}{\Longrightarrow}$$
$$C_2 = \begin{pmatrix} 382.5 & 117.8 & 0 & -48.8 \\ 402.2 & 63.8 & 0 & -26.4 \\ 127.5 & -117.8 & 0 & 48.8 \\ -28.6 & -153.9 & 0 & 63.8 \end{pmatrix} \quad (1)$$

$$E_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 255 & 255 & 0 \\ 0 & 0 & 0 & 0 & 255 & 255 & 0 & 0 \\ 0 & 0 & 0 & 255 & 255 & 0 & 0 & 0 \\ 0 & 0 & 255 & 255 & 0 & 0 & 0 & 0 \\ 0 & 255 & 255 & 0 & 0 & 0 & 0 & 0 \\ 255 & 255 & 0 & 0 & 0 & 0 & 0 & 0 \\ 255 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{after2}-\text{DDCT}}{\Longrightarrow}$$
$$C_3 = \begin{pmatrix} 414.4 & 125.9 & -100.5 & 66.2 & -31.9 & 5.9 & 7.2 & -7.5 \\ 125.9 & -250.6 & -246.2 & 83.9 & -58.2 & 2.9 & 7.2 & -12.2 \\ -100.5 & -246.2 & 80.8 & 271.2 & -17.3 & 31.0 & 22.5 & -6.5 \\ 66.2 & 83.9 & 271.2 & 106.1 & -204.7 & -29.5 & -7.1 & -21.5 \\ -31.9 & -58.2 & -17.3 & -204.7 & -223.1 & 86.7 & 41.7 & -6.0 \\ 5.9 & 2.9 & 31.0 & -29.5 & 86.7 & 228.1 & 16.0 & -25.0 \\ 7.2 & 7.2 & 22.5 & -7.1 & 41.7 & 16.0 & -144.6 & -51.3 \\ -7.5 & -12.2 & -6.5 & -21.5 & -6.0 & -25.0 & -51.3 & 43.9 \end{pmatrix}$$
$$E_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{after2}-\text{DDCT}}{\Longrightarrow}$$
$$C_4 = \begin{pmatrix} 414.4 & 71.3 & 41.7 & -78.1 & 31.9 & 2.1 & 17.3 & -31.8 \\ 273.0 & -20.1 & 49.0 & -55.3 & 37.5 & -22.0 & 20.3 & -16.7 \\ -287.3 & -174.9 & 22.5 & 32.7 & 17.3 & -49.0 & 9.3 & 25.3 \\ -462.7 & -78.6 & -11.5 & 42.2 & -8.8 & -14.4 & -4.8 & 20.5 \\ -159.4 & 159.7 & -41.7 & -3.1 & -31.9 & 52.2 & -17.3 & -14.2 \\ 176.1 & 197.2 & -57.8 & 4.3 & -44.2 & 66.6 & -23.9 & -14.8 \\ 222.5 & -4.8 & -54.4 & 70.9 & -41.7 & 18.1 & -22.5 & 23.9 \\ 98.0 & -125.5 & -32.7 & 85.7 & -25.0 & -17.9 & -13.6 & 38.8 \end{pmatrix} \quad (2)$$

II. ANALYSIS FOR INTRA RESIDUAL IN H.264/AVC

In H.264, intra prediction works well on the assumption that the samples change gradually along a certain direction in an intra prediction block. While in natural image, the boundary of an object is usually irregular and the change of samples can't be so homogeneous. As a result, intra residual still retains significant texture information comparing with the original frame as shown in Fig.2. So some work to adjust the intra residual to be suitable for 2D DCT will be very useful.

To illustrate the effectiveness of adjusting transform scheme for intra residual, some toy examples are shown below in (1)-(2) as mentioned in [6]. Here, equation (1) represents the sample of 4x4 transform coefficient energy distribution, and equation (2) gives the sample for 8x8. In (1)-(2), E1, E3 are the original intra residual matrices with directional edges, and E2, E4 are the reordered intra residual matrices with horizontal edges. Further, Fig.3 shows the energy distribution of the transform coefficients, where the x-axis denotes the sum of x and y components of the coefficients in the transform coefficient matrix, and the y-axis denotes the percentage of the energy of the transform coefficients. We can see that by reordering the direction of the edge to regular horizon, more energy in the transform domain concentrates in the low frequency part and this facilitates the entropy coding and improves the coding efficiency.

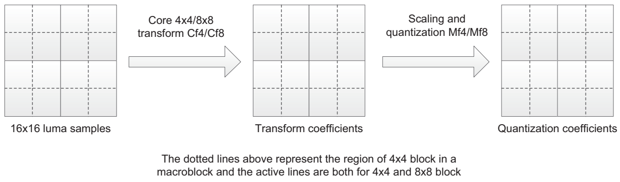


Fig. 4. The intra prediction mode directions of 4x4/8x8 and 16x16 block

On the other hand, to be suitable for limited-precision integer arithmetic implementation and to avoid the mismatch of encoder and decoder, integer core transform and normalizing steps with the quantization process have been employed in H.264 as shown in Fig.4. For the transform based on block, a simple method to guarantee the compatibility of our proposed algorithm is not to change the shape of the block. In this process, we just focus on adjusting core 4x4 transform and core 8x8 transform for simplifying the issue.

III. DESCRIPTION OF THE PROPOSED ALGORITHM

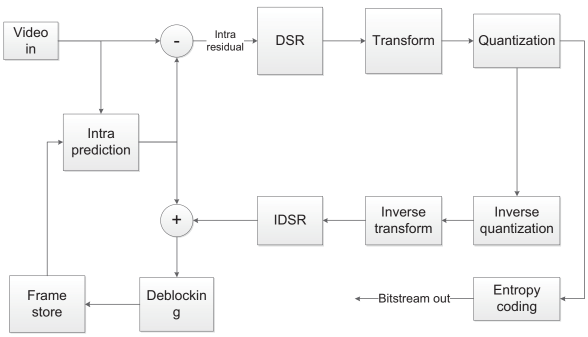


Fig. 5. Implement DSR in the H.264/AVC framework

Fig.5 shows the diagram of DSR in the H.264/AVC framework. Comparing with the original intra prediction process, we add two steps to improve the performance of transform as follows:

- Implement DSR for intra residual before transform and quantization process.
- Implement IDSR for the reordering intra residual to restore it to original order.

In the assumption that the intra residual in a block will carry on some characteristic associated with its intra prediction mode, we design eight scan modes to implement DSR. To make it clearly, the eight edge directions in intra residual and their corresponding scan mode have all been shown in Fig.6.

The angles of edge direction in Fig.6 are listed below (only acute angle is selected) :

- Vertical: $90°$ to horizontal
- Horizontal: $0°$ to horizontal
- Diagonal Down-Left: $45°$ to horizontal

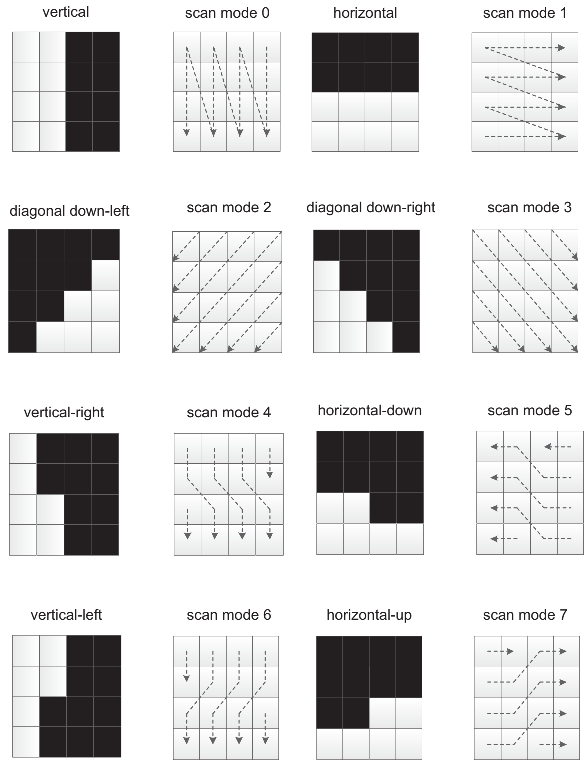


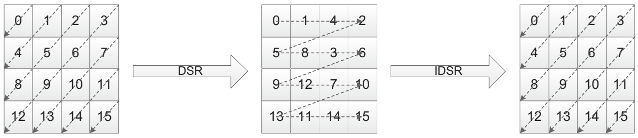
Fig. 6. Edge direction and scan mode for 4x4 intra prediction block

- Diagonal Down-Right: $45°$ to horizontal
- Vertical-Left: $26.6°$ to vertical
- Horizontal-Down: $26.6°$ to horizontal
- Vertical-Right: $26.6°$ to vertical
- Horizontal-Up: $26.6°$ to horizontal

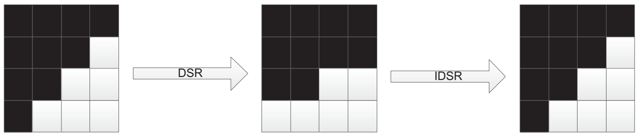
Here the eight edge directions shown in Fig.6 have the angles as similar as the 4x4 and 8x8 intra prediction modes mentioned above. And we can complete DSR procedure in the scan mode which can rotate the edge to the direction which is most close to horizon. Not similar with the other scan modes, we implement mirror flip in scan mode 0 and 1. Although this will not impact the energy distribution of transform coefficients for vertical and horizon edges, it changes coefficients' positive and negative nature which influences the run-length of coefficients at last. So we retain these two modes. To choose the most appropriate scan mode which can modify the residual's direction efficiently, a brute-force method is selected to implement DSR for intra residual blocks. And the scan mode which results in minimum rate-distortion cost (J) as the equation (3) shown will be chosen at last [11].

$$J = D + \lambda \times R \tag{3}$$

At the same time, extra three bits will be needed in the decoder to implement IDSR correctly for each encoded bock. Since the main purpose of this paper is to certify the efficiency of improving DCT performance by modifying residual block,



(a)



(b)

Fig. 7. DSR and IDSR implementing process in scan mode 2: this figure shows the residual positions' variances after DSR and IDSR process. In (a), the residual is marked by number. In (b), the residual is marked by color

the extra bits are not considered in our simulations for the time being. And the task of disposing the extra bits efficiently will also be left as one of our future works.

In each scan mode, we will take the samples according to its original order along the dotted line arrow shown in Fig.6, and put them into new position successively in raster order. Meanwhile, the inverse DSR (IDSR) is also needed to produce reconstructed image correctly. For example, we can describe the directional samples reordering procedure in scan mode 2 as shown in Fig.7, and other edge directions can be easily derived as the same way.

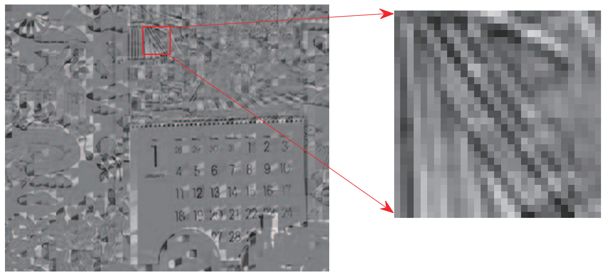
IV. EXPERIMENT RESULT

For simplifying the issue, we just implemented our DSR algorithm in 4x4 and 8x8 luma DCT procedures. The proposed method is implemented in the H.264/AVC reference software JM10.2 [12]. To enable 8x8 luma prediction, all experiments are conducted in H.264 High Profile. The key configuration of the encoder are listed in Table I. And the Bjontegaard Delta method [13] is used to evaluate the RD performance of our proposal DSR method compared with the high profile H.264 encoder.

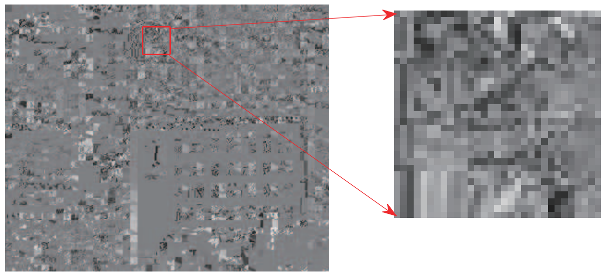
TABLE I
TEST CONFIGURATION

Item	Description of setting
Sequence	CIF from [14]
Frame number	150 frames each sequence
YUV format	4:2:0
Intra period	1(Only I frame is enabled)
Frame-rate	30fps
RDOptimization	1(High complexity mode)
QP	8,12,16,20
Entropy coding	CABAC

Fig.8 shows the subjective performance of mobile sequence for original intra residual and the reordering one. We can see that the texture in the region of calendar has been changed to significantly smooth and regular after DSR process.



(a) Diagonal edges in original residual



(b) Regular horizontal and vertical edges in reordered residual

Fig. 8. Comparison between original intra residual and reordered intra residual

The improvement of our proposal DSR method in PSNR compared with H.264 has been shown in Table II. On average, the PSNR gain is about 0.4372 dB, and for some sequences (e.g. *foreman*), up to 0.5967 dB gain can be achieved. What's more, at some points of high bitrate (e.g. 12000 kbps), the gain even can be close to 1 dB for foreman as shown in Fig.9 (a). And all these prove that our proposed method significantly outperforms H.264 in high bitrate region.

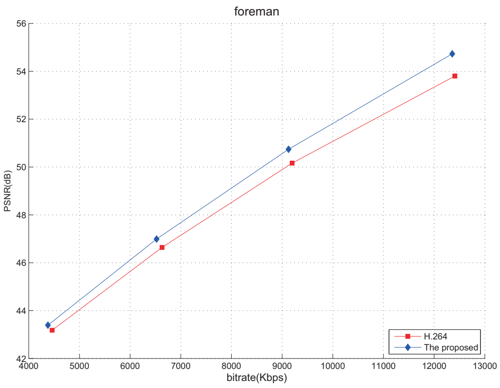
TABLE II
IMPROVEMNET OF THE PROPOSAL METOD

Sequence	BDPSNR(dB)	BDBR(%)
foreman	0.5967	-5.3558
mobile	0.5907	-4.3123
hall	0.3227	-3.1610
mother-daughter	0.2385	-2.8584
Average	0.4372	-3.9219

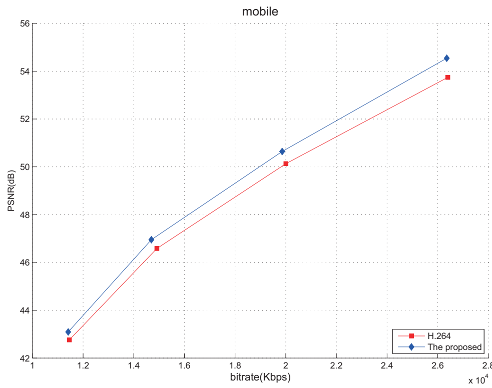
The RD-curves of all these test sequences in Table II have also been shown in Fig.9.

As shown in Table II and Fig.9, for some test sequences, such as hall and mother-daughter, the gain of our proposal DSR method is not as significant as for other sequences. To explain the reason of the fact, we analyze the MB types' distribution of all these test sequences in H.264 intra prediction process as shown in Fig.10.

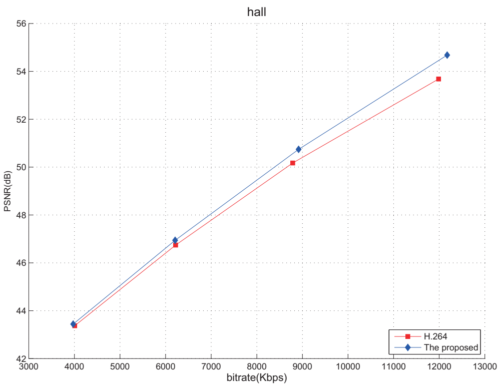
From Fig.10, we can see that for the sequences *hall* and *mother-daughter* with simple background, the 16x16 intra prediction block size usually takes majority in a frame. And as mentioned above, we just implemented our DSR algorithm in 4x4 and 8x8 luma DCT procedure for simplifying the issue. Of course, we can easily extend the DSR method to 16x16 block



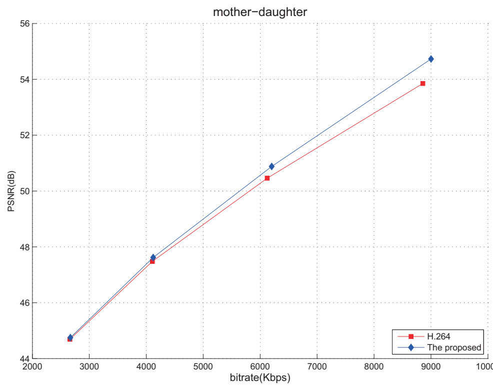
(a) RD-curve of foreman



(b) RD-curve of mobile



(c) RD-curve of hall



(d) RD-curve of mother-daughter

Fig. 9. Experimental results for CIF test sequences "foreman," "mobile," "hall," and "mother-daughter".

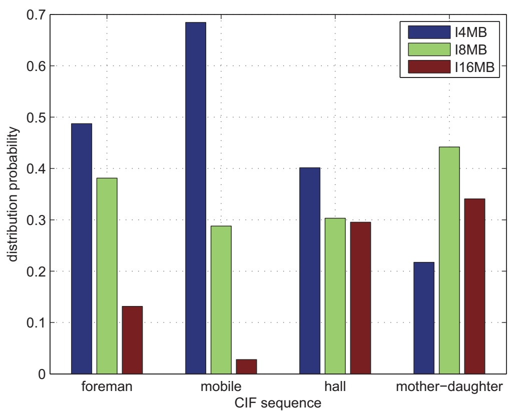


Fig. 10. MB types' distribution of test sequences "foreman," "mobile," "hall," "mother-daughter"

size as like it is in 4x4 and 8x8 block size, but the computation complexity will be too high to stand. On the other hand, there are some obvious differences for the luma forward transform procedure between intra 16x16 mode and 4x4, 8x8 mode in Fig.4.

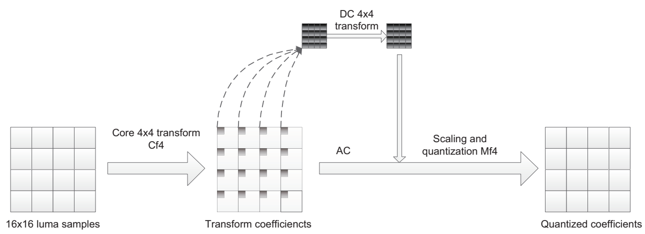


Fig. 11. Luma transform for intra 16x16 mode

As shown in Fig.11, macroblock predicted using intra 16x16 mode will need a second transform for the lowest or 'DC' frequency coefficients of the first transform [15], the impact of DSR to this DC transform is still under researching. What's more, since larger block size may contain more objects, this characteristic is also needed attention in our future work.

V. CONCLUSION

In this paper, we propose an efficient directional samples reordering algorithm for intra coding in H.264/AVC, in which the intra residual can be adjusted to a new order that is more suitable for 2D DCT. Experimental results show that it achieves better RD performance than H.264/AVC at high complexity mode. And now, there are still two issues in our future work. The first one is that a fast algorithm will be needed urgently, since the complexity is too high for every intra prediction block to traverse all these eight scan modes. The other one is that some improved DSR method for larger size intra prediction block need to be researched for HD sequence and low resolution sequence with simple background.

REFERENCES

[1] *Draft ITU-T Rec. and FDIS of Joint Video Spec. (H.264 — ISO/IEC 14496-10 AVC)*, JVT of MPEG and VCEG, Doc. JVT-G050rl, May 2003.

[2] G. B. T. Wiegand, G. Sullivan and A. Luthra, "Overview of the h.264/avc video coding standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, July 2003.

[3] L. L. Wang and W. C. Siu, "H.264 fast intra mode selection algorithm based on direction difference measure in the pixel domain," in *Proc. IEEE ICASSP '09*, Taiwan, April 2009, pp. 258–261.

[4] T. N. N. Ahmed and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Computer*, vol. 23, no. 1, pp. 90–93, January 1974.

[5] B. Zeng and J. Fu, "Directional discrete cosine transform - a new framework for image coding," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 18, no. 3, pp. 305–313, March 2008.

[6] C. Zhang, K. Ugur, J. Lainema, A. Hallapuro and M. Gabbouj, "Video Coding Using Spatially Varying Transform," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 21, no. 2, pp. 127–140, February 2011.

[7] P. L. D. V. Velisavljevic and M. Vetterli, "Directional wavelet transforms and frames," in *Proc. IEEE ICIP '02*, Rochester, NY, September 2002, pp. 589–592.

[8] M. V. V. Velisavljevic, B. Beferull-Lozano and P. L. Dragotti, "Discrete multi-directional wavelet bases," in *Proc. IEEE ICIP '03*, Barcelona, Spain, September 2003, pp. 1025–1028.

[9] M. N. Do and M. Vetterli, "Contourlets: A directional multiresolution image representation," in *Proc. IEEE ICIP '05*, Genova, Italy, September 2005, pp. 357–360.

[10] K. C. Hui and W. C. Siu, "New pixel-dct domain coding technique for object base and frame based prediction error," in *Proc. IEEE ICASSP '05*, Philadelphia, USA, March 2005, pp. 377–380.

[11] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag*, vol. 15, no. 6, pp. 74–90, November 1998.

[12] JVT H.264/AVC Reference Software version JM10.2, http://iphome.hhi.de/suehring/tml/download/old_jm/

[13] G. Bjontegaard, "Calculation of average psnr differences between rd-curves," Austin, Texas, USA, presented at the 13th VCEG-M33 Meeting, April 2001.

[14] Xiph.org test media, http://media.xiph.org/video/derf/

[15] I. E. Richardson, Ed., *The H.264 advanced video compression standard, Second Edition*. John Wiley & Sons. Ltd, UK, 2010.