7th WSEAS Int. Conf. on MATHEMATICAL METHODS and COMPUTATIONAL TECHNIQUES IN ELECTRICAL ENGINEERING, Sofia, 27-29/10/05 (pp235-238)

Scan Order and Huffman Coding of 3D DCT Coefficients

TOMÁŠ FRÝZA
Institute of Radio Electronics
Brno University of Technology
Purkyňova 118, 612 00 Brno
CZECH REPUBLIC

Abstract: - This paper is about video compression method based on the 3D DCT. Properties of this algorithm depend on the way how 3D DCT coefficients are grouped together into output bit stream. In the next text we briefly describe the main components of the encoder and decoder architecture including the definitions of mathematical operations of the Forward and Inverse Three Dimensional Discrete Cosine Transform. Secondly, two scan orders for 3D DCT coefficients are presented. Also the modified Huffman dictionary for 3D DCT coefficients is discussed. The experimental confirmation of proposed methods is realized as well.

Key-Words: - Video compression, 3D DCT, Scan order, Huffman coding, Compress ratio, NRMSE

1 Introduction

The main purpose of all video methods is focused on a reduction of correlation between the image points inside one frame (so-called intraframe coding), likewise the correlation between the image points in the adjacent frames (intraframe coding). This paper is concentrate on the video compression techniques with the 3D DCT (Three Dimensional Discrete Cosine Transform). This transform coding has an advantage that join together intraframe and interframe codings. Although the 3D DCT can reached the decorrelation of input pictures elements, the number of calculated coefficients stays unchanged. The bitrate reduction is achieved after entropy coding, in relation to scan order and Huffman codewords for 3D DCT coefficients. The main objective of this article is to describe two methods of 3D DCT coefficients scan orders and the useful Huffman codewords dictionary.

The paper is divided into three major parts. In the first part, *Section 2*, the brief description of the 3D DCT encoder is presented. The second part, *Section 3*, is devoted to the scan order for 3D DCT coefficients and *Section 4* presents proposed Huffman codewords. Comments and results are given in the last section (*5*).

2 3D DCT Transform

The Forward Three Dimensional Discrete Cosine Transform (3D fDCT) for one color palette and for group of $N \times N \times N$ picture elements is given by [5]

$$X^{C(2)}(l,m,n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{f=0}^{N-1} x(i,j,f) \cdot \left[C_N^{\mathrm{II}}\right]_{li} \cdot \left[C_N^{\mathrm{II}}\right]_{mj} \cdot \left[C_N^{\mathrm{II}}\right]_{nf} \qquad (1)$$

where $x(i,j,f)$ is an input sequence pixel's value at position of (i,j,f), $X^{C(2)}(l,m,n)$ represents its transformed value, $l = m = n = 0,1,\ldots,N-1$ and $\left[C_N^{\mathrm{II}}\right]$ describes the 1D DCT base function defined as [2]

$$\left[C_N^{\mathrm{II}}\right]_{prm0\ prm1} = \left(\frac{2}{N}\right)^{1/2} \left[\gamma_{prm0}\ \cos\left(\frac{prm0\ (prm1+\frac{1}{2})\ \pi}{N}\right)\right] \qquad (2)$$

where $prm0,\ prm1 = 0,1,\ldots,N-1$ are input parameters. The weighting functions γ_{prm0} and γ_{prm1} are depicted by equation [5]

$$\gamma_{prm0},\ \gamma_{prm1} = \begin{cases} \frac{1}{\sqrt{2}} & \text{if} \quad prm0,\ prm1 = 0 \\ 1 & \text{if} \quad prm0,\ prm1 \neq 0. \end{cases} \qquad (3)$$

The next part of the 3D DCT encoder is the process when the insignificant coefficients are suppressed. The consequence is the reduced number of non-zero coefficients and therefore increase the compress ratio. This aim is called quantization. The proces of video sequence encoding is followed by therholding, where the coefficients smaller than the threshold value are abolished. And finally, last part of the encoder is an entropy coding. This paper tries to improve the compression

Research described in the paper was financially supported by the Czech Grant Agency under grants No. 102/03/H109 and by the research program MSM 021630513.

properties by using of adequate scan order and Huffman codewords optimized for the 3D DCT coefficients.

On the decoder size, the authentic coefficients values are reconstructed by dequantization operation and last part of the decoder is the Inverse Three-Dimensional Discrete Cosine Transform (3D iDCT) [5]

$$\hat{x}(i,j,f) = \sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} \hat{X}^{C(2)}(l,m,n) \cdot \left[C_N^{\mathrm{II}}\right]_{li} \cdot \left[C_N^{\mathrm{II}}\right]_{mj} \cdot \left[C_N^{\mathrm{II}}\right]_{nf} \tag{4}$$

where $\hat{x}(i,j,f)$ is reconstructed picture element at position of (i,j,f) and $\hat{X}^{C(2)}(l,m,n)$ is frequency coefficients after dequantization.

3 Scanning order of 3D DCT coefficients

After the transform coding the picture elements are transformed into frequency coefficients. The quantized coefficient $X_Q^{C(2)}(0,0,0)$ is called DC and represents the total luminance in encoded block. The remaining $N \times N \times N - 1$ coefficients are called AC coefficients. Both, DC and AC coefficients are encoded separately. Forming these coefficients into output bit stream is provided by the entropy coding, which is lossless operation. Therefore, the error of the reconstructed video sequence cannot be increased by this operation. On the other hand, the compress ratio depends on the way how the coefficients are regrouped and on the Huffman coding capacity. In the next text, two ways of 3D DCT coefficients scan orders are presented.

3.1 Modified zig-zag method

First method, how 3D DCT coefficients can be reordered is called Modified zig-zag method and it was derived from scan order of 2D DCT coefficients in JPEG (Joint Picture Experts Group) compression standard [4]. All $N \times N \times N$ coefficients are divided into N planes and the coefficients $X_Q^{C(2)}(l,m,n)$ inside each plane $n = 0,1,\ldots,N-1$ are scanned in original zig-zag way. This procedure is displayed in *Figure 1*.

The purpose of reordering of the 3D DCT coefficients is to achieved the gradual descent of the AC values. In this situation, the coefficients can be effectively encoded by Huffman coding. The proposed method was applied on two test video sequences. The AC coefficient values from both of them are depicted in *Figure 2*. First real sequence (titled `sky`) contains low number of details in pictures and slow motion. This causes small number of non-zero coefficients in frequency domain. Second sequence (`competition`)

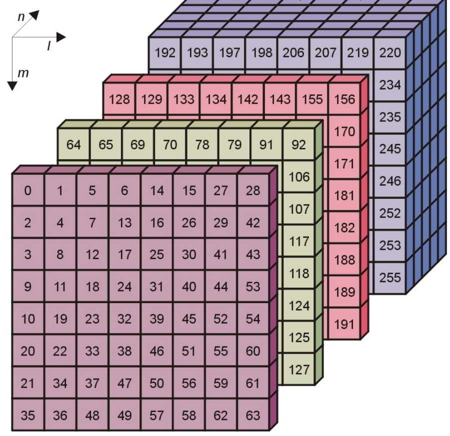
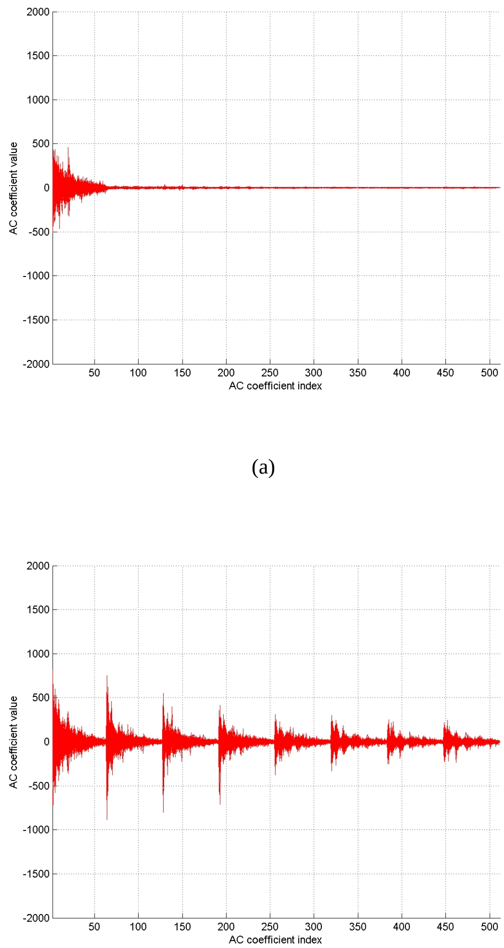


Figure 1: Scan order for 3D DCT coefficients by modified zig-zag method.



(a)

(b)

Figure 2: AC coefficients ordered by modified zig-zag method from sequence: a) `sky` and b) `competition`.

represents entirely different type of scene. The considerable number of AC coefficients corresponds with fast motion and high details in pictures. The presents of outstanding peaks in displayed series of AC coefficients implies to non-optimal scan order.

3.2 3D DCT optimized method

Second proposed method for regrouping the AC coefficients corresponds with importance of particular 3D DCT coefficients. The coefficients with high values are distributed along the main axis in cube of transformed picture elements and for low values of indexes *l*, *m* and *n*. This property was took into account in second method of scan order. The cube of 3D DCT coefficients are separated into planes perpendicular with main diagonal in the cube. In these planes, the frequency coefficients are scanned in zig-zag form. This idea is shown in *Figure 3*.

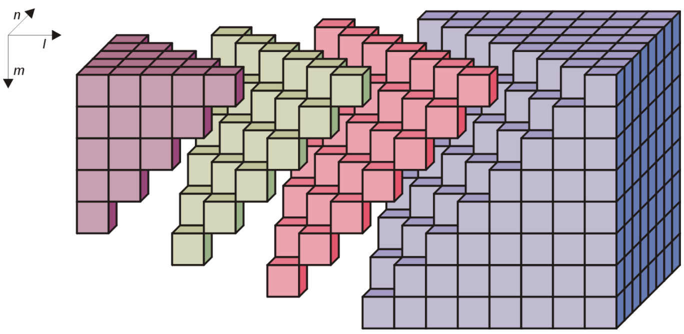


Figure 3: Scan order for 3D DCT coefficients by 3D DCT optimized method.

Analysis of proposed method with test sequences is depicted in *Figure 4*. Although the succession of nonzero AC coefficients in first sequence is disturbed by zero values, this method is preferable for reordering the 3D DCT coefficients. That proclaims the curve of AC coefficients of second video sequence. Here, no marked peaks in succession of AC coefficients are visible.

4 Huffman coding

The last part of 3D DCT encoder is entropy coding of reordered coefficients, realized by Huffman code. As mentioned above, the DC and AC values are encoded separately with different Huffman dictionaries. In addition, the DC coefficients are encoded with differential coding and AC coefficient are grouped to the symbols describing not only the actual AC value but also the number of precedents coefficients with zero values (so-called run length).

For requisites of 3D DCT coefficients, two Huffman dictionaries were derived (for DC and AC coefficients). The concrete codewords were found out by real sequences analysis. In *Table 1* the length of each

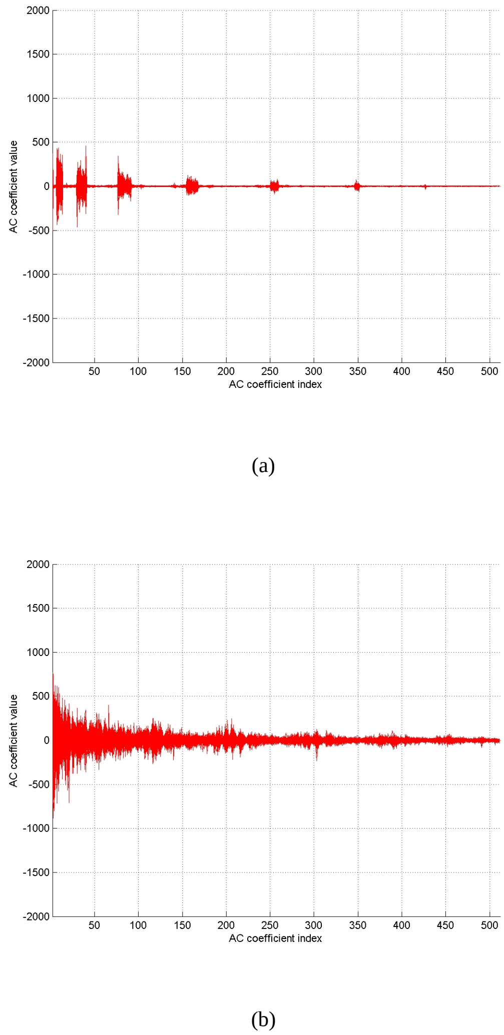


Figure 4: AC coefficients ordered by 3D DCT optimized method from sequence: a) `sky` and b) `competition`.

code words in bits for DC coefficients are presented. The lowest values of bits were reached for DC coefficients from interval $\langle -1; 31 \rangle$. The values with smaller frequency are encoded with longer codewords.

The codewords length along with run length for AC coefficients are shown in *Table 2*, where symbol EOB represents the end-of-block identifier. Every encoded block of transformed picture elements is terminated by this symbol. From mentioned results are obvious that the most frequent combinations of run length and AC coefficient values are coefficients -1 and 1.

Also smaller values of AC coefficients are more frequent in every group of symbols with constant value of run length. At first sight there would be indicated that phenomenal length of codewords for symbol group with maximum value of run length, cannot ensure useful compress ratio. But in this situation the single codeword can describe all AC coefficients in transformed

Table 1: Codewords lengths for DC coefficients

Run Length	*DC coefficient value*		*Codeword length [b]*
–	0		4
–	1(-1)	1(-1)	3
–	2(-3)	3(-2)	2
–	4(-7)	7(-4)	3
–	8(-15)	15(-8)	3
–	16(-31)	31(-16)	3
–	32(-63)	63(-32)	4
–	64(-127)	127(-64)	4
–	128(-255)	255(-128)	5
–	256(-511)	511(-256)	6
–	512(-1.023)	1.023(-512)	7
–	1.024(-2.047)	2.047(-1.024)	8
–	2.048(-4.095)	4.095(-2.048)	8

Table 2: Codewords lengths for AC coefficients

Run Length	*AC coefficient value*		*Codeword length [b]*
0	EOB		6
0	1(-1)	1(-1)	2
0	2(-3)	3(-2)	3
0	4(-7)	7(-4)	3
0	8(-15)	15(-8)	4
0	16(-31)	31(-16)	5
0	32(-63)	63(-32)	7
0	64(-127)	127(-64)	8
0	128(-255)	255(-128)	11
0	256(-511)	511(-256)	15
0	512(-1.023)	1.023(-512)	26
0	1.024(-2.047)	2.047(-1.024)	27
1	1(-1)	1(-1)	3
1	2(-3)	3(-2)	5
1	4(-7)	7(-4)	7
1	8(-15)	15(-8)	8
1	16(-31)	31(-16)	10
1	32(-63)	63(-32)	12
1	64(-127)	127(-64)	14
1	128(-255)	255(-128)	17
1	256(-511)	511(-256)	20
1	512(-1.023)	1.023(-512)	28
1	1.024(-2.047)	2.047(-1.024)	29
⋮	⋮		⋮
510	1(-1)	1(-1)	5.080
510	2(-3)	3(-2)	5.081
510	4(-7)	7(-4)	5.082
510	8(-15)	15(-8)	5.083
510	16(-31)	31(-16)	5.084
510	32(-63)	63(-32)	5.085
510	64(-127)	127(-64)	5.086
510	128(-255)	255(-128)	5.087
510	256(-511)	511(-256)	5.088
510	512(-1.023)	1.023(-512)	5.089
510	1.024(-2.047)	2.047(-1.024)	5.089

video cube. The maximum length codewords results from number of all items in Huffman dictionary. Let length od Discrete Cosine Transform is $N = 8$, then the number of all elements in AC coefficient dictionary is 5.622.

5 Conclusion

The consequences of proposed scan orders and Huffman dictionaries to compression properties were tested on a set of seven real video sequences. Two compression properties were observed: the CR (Compress Ratio) defined by [1] and the quality degradation in output sequence, evaluated by NRMSE (Normalized Root Mean Square Error) [3]. Results are shown in *Table 3*. It is obvious that scan order derived from 3D DCT coefficients values can reached higher compress ratio in contrast to the standard zig-zag scanning. With all test sequences and for two required picture qualities the scan order optimized for 3D DCT coefficients can guarantee higher compress ratio with keeping the same error in output video sequences.

Table 3: Compress properties in dependence on the 3D DCT coefficient scan orders

Sequence name	*CR (high quality)*		*NRMSE*
	zig-zag	*3d dct*	
competition	12,69	13,42	0,0353
high_jump	17,45	18,58	0,0363
school	24,14	24,63	0,0255
sky	51,40	54,67	0,0129
street0	17,61	18,54	0,0310
street1	18,46	19,35	0,0267
wind	6,25	6,47	0,0581

Sequence name	*CR (middle quality)*		*NRMSE*
	zig-zag	*3d dct*	
competition	28,49	30,77	0,0566
high_jump	44,88	48,07	0,0623
school	59,92	59,63	0,0442
sky	145,48	140,76	0,0210
street0	41,98	44,47	0,0537
street1	42,35	45,51	0,0497
wind	12,76	13,46	0,1096

Reference:

[1] T. Fryza, S. Hanus, Relation Between Character of Real Video Sequences and 3D DCT Compression, In *Proceedings EC-VIP-MC 2003*, FEEC Zagreb: Zagreb (Croatia) 2003, pp. 107-112.

[2] K.R. Rao, P. Yip, *Discrete Cosine Transform. Algorithms, Advantages, applications*, Academic Press, Inc., San Diego (USA) 1990.

[3] E. Saavedra, A. Grauel, D. Morton, Combined Methods for Image Compression, In *Recent Advances in Intelligent Systems and Signal Processing*, WSEAS: Corfu (Greece) 2003, pp. 233-235.

[4] G.K. Wallace, The JPEG Still Picture Compression Standard, *IEEE Transactions on Consumer Electronics*, Vol.38, No.1, 1992, pp. xviii-xxxiv.

[5] R. Westwater, B. Furht, *Real-Time Video Compression. Techniques and Algorithms*, Kluwer Academic Publishers, Boston (USA) 1997.