

Shape-adaptive DCT and Its Application in Region-based Image Coding

Yamin Zheng, Xiaoyan Wang and Chengyou Wang

*School of Mechanical, Electrical and Information Engineering, Shandong University,
Weihai 264209, P. R. China*

sduzym@vip.sina.com, swwxy00800313@163.com, wangchengyou@sdu.edu.cn

Abstract

To get better and flexible performance, region-based image coding has been studied widely in the last few years. In region-based image coding, the whole image can be divided into two parts: region-of-interest (ROI) and background area. ROI is coded in low compression ratio, while background area is coded in high compression ratio. In this paper, shape-adaptive DCT (SA-DCT) for coding arbitrarily shaped image segments is introduced. And two region-based image coding methods using discrete cosine transform (DCT) and SA-DCT are presented. The algorithm based on SA-DCT is comparable with the one based on DCT in terms of reconstructed images' effects. Experimental results show that the algorithm based on SA-DCT performs better than the one based on DCT both in subjective effect and objective effect. Therefore, SA-DCT is a more suitable transform for region-based image coding than DCT.

Keywords: *image coding; discrete cosine transform (DCT); shape-adaptive DCT (SA-DCT); region-of-interest (ROI); image segmentation*

1. Introduction

Conventional image coding algorithm which uniformly partitions an image into $N \times N$ rectangle blocks has some problems, such as mosquito and blocking artifacts [1], especially in very low bit rate applications. To solve this problem, region-based image coding has been studied widely in the last few years. Region-based image coding consists of three steps: image segmentation, contour coding and texture coding. Texture coding is mainly introduced in this paper. An image is segmented into region-of-interest (ROI) and background area by image segmentation [2]. Encoder can encode ROI and background area respectively. Therefore decoder can decode and manipulate ROI independently from the background area. And region-based image coding can improve image coding quality in low bit rate applications.

Shape-adaptive DCT (SA-DCT) is a transformation method for coding arbitrarily shaped image segments, and image denoising [1, 3]. The region-based image coding includes region-of-interest and background area coding. In this paper, two region-based image coding methods using discrete cosine transform (DCT) and SA-DCT are achieved. The region-based image coding algorithm using SA-DCT uses Δ DC-SA-DCT transform method [4]. The region-based image coding method using DCT uses DCT-JPEG coding system framework [5]. Compared with region-based image coding method using DCT, the region-based image coding using SA-DCT performs better both in subjective effect and objective effect.

The rest of this paper is organized as follows. Section 2 introduces DCT and SA-DCT. And then in Section 3, the region-based image coding is explained. Experimental results of

the region-based image coding using DCT and SA-DCT are given in Section 4. Some reconstructed images are presented. Conclusions and remarks on possible further work are given finally in Section 5.

2. DCT and Shape-adaptive DCT

2.1. Discrete Cosine Transform (DCT)

The traditional two-dimensional DCT transform is made usually by two one-dimensional DCT transform on row and column directions separately. Let X and C represent an image block and DCT matrix with size of $N \times N$ respectively. After two-dimensional DCT transform, transform coefficients block Y can be denoted by

$$Y = CXC^T, \quad (1)$$

$$C(i, j) = \begin{cases} \sqrt{\frac{1}{N}}, & i = 0, j = 0, 1, \dots, N-1, \\ \sqrt{\frac{2}{N}} \cos \frac{i(2j+1)\pi}{2N}, & i = 1, 2, \dots, N-1, j = 0, 1, \dots, N-1. \end{cases} \quad (2)$$

where C^T is the transpose matrix of C .

2.2. DCT-based Image Coding

The simplest DCT-based image coding process is referred to as the baseline sequential process. Figure 1 shows the main procedures for encoding and decoding processes based on the conventional DCT [5]. It illustrates the special case of a single-component image; for color images, all processes operate on each image component independently.

In the encoding process, the input image data are grouped into 8×8 blocks, and each block is transformed by the forward DCT (FDCT) into 64 DCT coefficients [6]. One is the DC coefficient and the other 63 are the AC coefficients. Each of the 64 coefficients is then quantized using one of 64 corresponding values from the quantization table. The JPEG standard [5] suggests the luminance and chrominance quantization table respectively, where the size of quantization interval is determined by lots of subjective experiments.

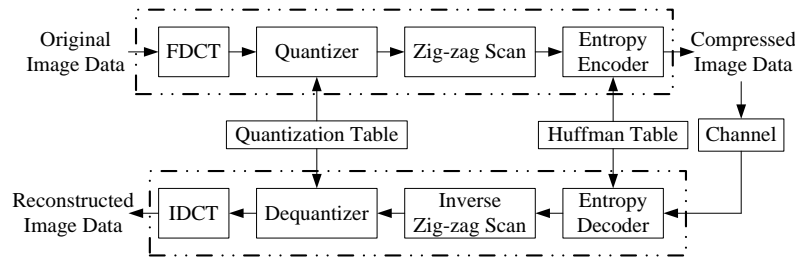


Figure 1. Simplified Diagram of DCT-based Image Encoder and Decoder

After quantization, the DCT coefficients are prepared for entropy encoding. The previous quantized DC coefficient is used to predict the current quantized DC coefficient, and the difference is encoded. The 63 quantized AC coefficients undergo no such differential encoding, but they are converted into a one-dimensional zig-zag sequence. The quantized

coefficients are then passed to an entropy encoding procedure which compresses the data further. Here, Huffman encoding is used and Huffman table must be provided to the encoder.

In contrast to the encoder, each step of decoding process performs essentially the inverse of its corresponding main procedure within the encoder. The entropy decoder decodes the zig-zag sequence of quantized DCT coefficients. After dequantization the DCT coefficients are transformed to an 8×8 block of samples by the inverse DCT (IDCT).

2.3. Shape-adaptive DCT (SA-DCT)

The process of SA-DCT is composed of four steps: vertical shift, vertical one-dimensional variable-length DCT, horizontal shift and horizontal one-dimensional variable-length DCT. The SA-DCT is defined as follows [4]. Figure 2(a) shows a boundary block of an arbitrarily shaped image object segmented into foreground (black) and background (white). To perform SA-DCT to foreground segment, firstly the position of first pixel of each column in the foreground and the length of each column in the foreground are obtained. Then based on the obtained shape information, all pixels of each column in foreground segment are shifted to the uppermost position and grouped into column vectors \mathbf{x}_j (Figure 2(b)). Then, N_j pixels of each column vector \mathbf{x}_j are transformed by vertical one-dimensional DCT into N_j vertical transform coefficients (Figure 2(c)). The first coefficient of each transform coefficients column vector \mathbf{a}_j is DC coefficient. Based on the position of first coefficient of each row and length of each row in the foreground, all transform coefficients of each row are shifted to the left border and grouped into row vectors \mathbf{b}_i (Figure 2(d)). M_i transform coefficients of each row in the foreground segment are transformed by horizontal one-dimensional DCT into final SA-DCT coefficients (Figure 2(e)) [1, 2, 7]. The vertical and horizontal one-dimensional DCT transforms are expressed in (3) and (4).

$$\mathbf{a}_j = S_{N_j} \times DCT_{N_j} \cdot \mathbf{x}_j, \quad (3)$$

$$\mathbf{c}_i = S_{M_i} \times DCT_{M_i} \cdot \mathbf{b}_i, \quad (4)$$

where DCT_L ($L = N_j$ or $L = M_i$) denotes the one-dimensional DCT transform matrix with size of $L \times L$ defined in (5) and S_L ($L = N_j$ or $L = M_i$) is the SA-DCT transform prefactor. The horizontal and vertical inverse one-dimensional DCT transforms are expressed in (6) and (7).

$$DCT_L(p, k) = c_0 \cos\left(p\left(k + \frac{1}{2}\right)\frac{\pi}{L}\right), \quad c_0 = \begin{cases} \sqrt{1/2}, & p=0, \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

$$\mathbf{b}_i^* = \frac{2}{M_i S_{M_i}} \times DCT_{M_i}^T \cdot \mathbf{c}_i^*, \quad (6)$$

$$\mathbf{x}_j^* = \frac{2}{N_j S_{N_j}} \times DCT_{N_j}^T \cdot \mathbf{a}_j^*, \quad (7)$$

where the stars indicate the resulted data after the SA-DCT coefficients are quantized and dequantized. The $DCT_{M_i}^T$ and $DCT_{N_j}^T$ denote the transpose matrix of DCT_{M_i} and DCT_{N_j} .

To avoid mean weighting defect, the scaling factors S_L should be inverse proportional to L . Let $S_L = 4/L$, however, this version of SA-DCT is not normalized, which leads to low coding efficiency and noise weighting defect [4]. To avoid noise weighting defect, the scaling factors S_L should be set to $\sqrt{2/L}$. This kind of SA-DCT is called pseudo-orthonormal SA-DCT (PO-SA-DCT) [4]. To solve mean weighting defect and noise weighting defect, Δ DC-SA-DCT transform method is optimal. The corresponding block diagram is shown in Figure 3.

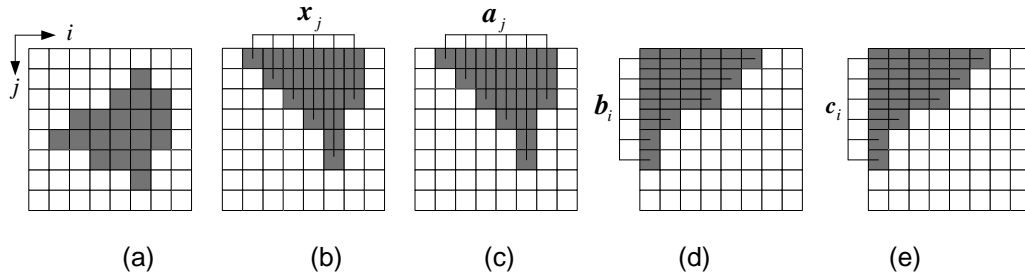


Figure 2. Forward SA-DCT of Foreground Segment of a Boundary Block: (a) Original Boundary Block, (b) The Boundary Block after the Pixels in the Foreground Segment are Shifted Uppermost Position, (c) The Boundary Block after Vertical One-dimensional DCT, (d) The Boundary Block after the Foreground Segment Coefficients are Shifted Leftmost Position, (e) The Final SA-DCT Transform Coefficients Boundary Block after Horizontal one-dimensional DCT

Before PO-SA-DCT, the encoder computes the average m of the input foreground segment image data. Then the average m is subtracted by the input foreground segment image data. The residual data with mean value 0 are transformed by PO-SA-DCT into SA-DCT coefficients. The mean value is scaled in the same way as DC values in DCT. The SA-DCT coefficient $c_{1,1}$ is overwritten with the scaled mean value DC.

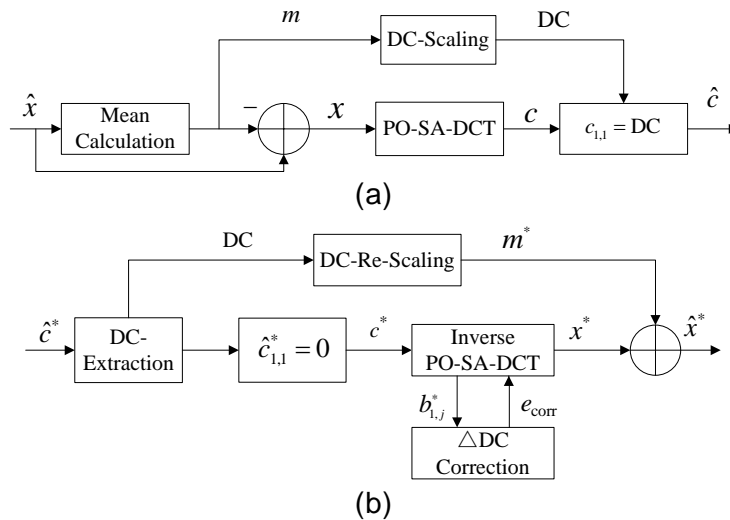


Figure 3. Block Diagram of Δ DC-SA-DCT: (a) Forward Transform, (b) Inverse Transform

The decoder extracts the DC coefficient from the received coefficients and the DC coefficient is rescaled to mean value m^* . The DC coefficient $\hat{c}_{1,1}^*$ is set to zero. Each row vector c_i^* of the resulted PO-SA-DCT coefficients is transformed by horizontal inverse one-dimensional DCT into b_i^* . Then the corrected value e_{corr} is obtained by

$$e_{\text{corr}} = \frac{\sum_{j=1}^M \sqrt{N_j} b_{1,j}^*}{\sum_{j=1}^M \sqrt{N_j}}, \quad (8)$$

where $b_{1,j}^*$ denotes each coefficient of first row vector b_1^* and M is the length of b_1^* ; N_j is the length of each column of original foreground segment; e_{corr} is subtracted by $b_{1,j}^*$ to correct the deviation. Then each row element of b_i^* is shifted to original right position according to block shape information and grouped into column vector a_j^* . Each column vector a_j^* is transformed by vertical inverse one-dimensional DCT into x_j^* and each element of x_j^* is shifted to original position towards down according to block shape information. Finally, m^* is added to shifted $x_{i,j}^*$ and the reconstructed block image data are got.

3. Region-based Image Coding

Region-based image coding consists of three steps: image segmentation, contour coding, and texture coding. An image is manually split into the ROI and background area. Contour coding is not studied in this paper. Texture coding includes ROI and background area texture coding. The ROI is coded independent of background area. Figure 4 shows the simplified diagram of encoding and decoding system based on SA-DCT.

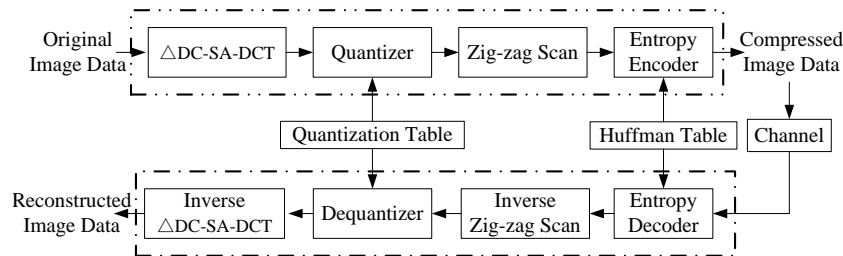


Figure 4. Simplified Diagram of Encoding and Decoding System based on SA-DCT

Background area image data are set to zero, when the ROI image data are encoded [8]. A minimum rectangle including region-of-interest is found, whose length and width are multiples of eight. The upper-left pixel's row and column number in the rectangle are multiples of eight plus one.

Region-of-interest encoding and decoding processes based on SA-DCT are described below. The pixels in minimum rectangle including region-of-interest are grouped into 8×8 blocks, and the number of non-zero pixels of each block is calculated.

Based on the calculated number of non-zero pixels of each block, all blocks are divided into three categories:

- (a) The block with 0 non-zero pixels is not transformed.

- (b) The block with 64 non-zero pixels is transformed by standard DCT.
- (c) The block with non-zero pixels number is larger than 0 and smaller than 64 is transformed by Δ DC-SA-DCT [4].

Each of the resulted coefficients is quantized using the corresponding value in the quantization table. And then the quantized data are entropy coded.

In the decoding process, the transmitted data are entropy decoded. Then the decoded data are dequantized. According to the three different categories of blocks, different transforms are performed.

- (a) The dequantized data of the block with 0 non-zero pixels are not inverse transformed.
- (b) The dequantized data of the block with 64 non-zero pixels are transformed by inverse DCT.
- (c) The dequantized data of other blocks are transformed by inverse Δ DC-SA-DCT transform method.

Now the reconstructed image data of region-of-interest are obtained.

Region-of-interest image data are set to zero, when background area image data are encoded. The encoding and decoding processes of background area image data are the same as the region-of-interest. Finally the reconstructed region-of-interest and background area image data are merged to get the reconstructed full image.

4. Experimental Results

4.1. Region-based Image Coding Using MATLAB

Simulation experiments are conducted using MATLAB 7.8. In order to highlight the advantages of region-based image coding using SA-DCT, we have compared the region-based image coding scheme using SA-DCT with the one based on DCT in terms of objective effect and subjective effect. Announcer and Lena (8bits/pixel, 512×512) are the test gray images. The region-of-interest images of both images with background area image data set to zero are shown in Figure 5. The contour data of the region-of-interest are not encoded in this paper. Experimental results are obtained in the case of the decoder has known contour data of the region-of-interest of both images.



**Figure 5. Region-of-interest with Background Area Image Data Set to zero:
(a) Announcer, (b) Lena**

In the case of SA-DCT, region-of-interest and background area are encoded separately according to the encoding process based on SA-DCT. The region-of-interest is encoded in low compression ratio, while the background area is encoded in high compression ratio. Then the encoding bit rates of region-of-interest and background area are calculated independently. Then both areas are decoded according to the decoding process based on SA-DCT. The PSNRs of both areas are computed by the reconstructed ROI and background area image data. Afterwards the reconstructed region-of-interest and background area image data are merged to obtain the full image and the PSNR of the reconstructed full image is computed.

In the case of DCT, the minimum rectangle including the region-of-interest with background area image data are set to zero is grouped into 8×8 blocks and each block is encoded according to the encoding process of DCT-JPEG. The encoding bit rate of ROI is calculated. Then the full image data with the region-of-interest image data are set to zero (i.e., background area) are encoded. The encoding bit rate of the background area image data is calculated. Afterwards the encoded region-of-interest image bits are decoded to get the reconstructed region-of-interest image data according to the decoding process of DCT-JPEG. The PSNR is computed by reconstructed region-of-interest image data. Then the encoded background area image data are decoded according to decoding process of DCT-JPEG. The PSNR is computed by reconstructed background area image data. Finally the reconstructed region-of-interest and background area image data are merged to obtain the reconstructed full image. Then the PSNR of the reconstructed full image is computed.

4.2. Comparison with DCT

Region-based image coding using SA-DCT is compared with the one based on DCT in terms of PSNRs of the reconstructed images including region-of-interest, background area and full image. The experimental results of the Announcer image using DCT and SA-DCT respectively are shown in Table 1.

Table 1. Comparison of Region-based Image Coding using DCT and SA-DCT Applied to Image Announcer

ROI bit rate (bpp)	Background bit rate (bpp)	PSNR (dB)					
		DCT			SA-DCT		
		ROI	Background	Full image	ROI	Background	Full image
0.60	0.30	35.75	33.72	34.42	37.85	34.42	35.49
0.70	0.40	36.69	35.46	35.91	38.83	37.06	37.69
0.80	0.50	37.53	36.84	37.10	39.58	38.35	38.80
0.90	0.60	38.28	37.78	37.98	40.33	39.18	39.60
1.00	0.70	38.93	38.57	38.71	40.89	39.95	40.31
1.10	0.80	39.55	39.23	39.35	41.40	40.53	40.86

As shown in Table 1, the region-of-interest encoding bit rate and the background area encoding bit rate based on SA-DCT are the same as DCT. The region-of-interest is compressed with low compression ratio, while the background area is compressed with high compression ratio. Experimental results show that the PSNRs of the reconstructed images (region-of-interest, background area, and full image) based on SA-DCT are higher than the ones based on DCT at various bit rates. Therefore, the region-based image coding algorithm using SA-DCT is better than the one based on DCT in terms of objective effect.

To compare the compression performance subjectively, Figure 6 and Figure 7 show the reconstructed images Announcer and Lena based on DCT and SA-DCT respectively in the case when ROI bit rate is 0.80bpp and background bit rate is 0.50bpp. According to the subjective effect, the reconstructed images including Announcer and Lena based on SA-DCT are better. The reconstructed images including Announcer and Lena based on DCT are worse than that based on SA-DCT especially in terms of the edge of the region-of-interest.

From the experimental results, we can see that compared to the algorithm based on DCT, the region-based image coding algorithm using SA-DCT improves compression performance in term of PSNR and visual quality. And the region-based image coding methods using SA-DCT and using DCT are also be applied to other test images, similar results can be obtained.



Figure 6. Reconstructed Announcer Images: (a) DCT, (b) SA-DCT



Figure 7. Reconstructed Lena Images: (a) DCT, (b) SA-DCT

5. Conclusion

In this paper, the transform method SA-DCT for coding arbitrarily shaped image segments is introduced. The region-based image coding process of an image consists of three steps: image segmentation, contour coding and texture coding. The texture coding includes ROI texture coding and background area texture coding. To reduce the total encoding bit rate and take advantage of the vision characteristics of human eyes, the region-of-interest is compressed in low compression ratio and the background area is compressed in high

compression ratio. Two region-based image coding methods using DCT and SA-DCT are achieved. Compared with region-based image coding method using DCT, the region-based image coding algorithm using SA-DCT is better both in terms of objective effect and subjective effect. Therefore, SA-DCT is more suitable for region-based image coding than DCT.

Texture coding is mainly introduced in this paper, while contour coding is not considered. So the contour coding is going to be studied in the further work.

Acknowledgements

This work was supported by the teaching reform project of Shandong University, Weihai, China (Grant No. B201319) and the National Natural Science Foundation of China (Grant No. 61201371). The authors wish to thank Songzhao Xie and Chunxiao Zhang for their help and valuable suggestions.

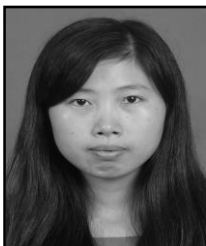
References

- [1] T. Sikora, "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments", *Signal Processing: Image Communication*, vol. 7, no. 4-6, (1995), pp. 381-395.
- [2] P. Garstecki and R. Stasinski, "Influence of image segment size on efficiency of SA DCT", *Proceedings of the 7th Nordic Signal Processing Symposium, Reykjavik, Iceland, (2006) June 7-9*, pp. 98-101.
- [3] A. Foi, K. Dabov, V. Katkovnik and K. Egiazarian, "Shape-adaptive DCT for denoising and image reconstruction", *Proceedings of SPIE - Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning, San Jose, CA, USA, vol. 6064, (2006) January 16-18*, pp. 0N1-0N12.
- [4] P. Kauff and K. Schueuer, "Shape-adaptive DCT with block-based DC separation and Δ DC correction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, (1998), pp. 237-242.
- [5] ISO/IEC, "Information Technology-Digital Compression and Coding of Continuous-tone Still Images-Part 1: Requirements and Guidelines", *ISO/IEC 10918-1, (1994)*.
- [6] G. K. Wallace, "The JPEG still picture compression standard", *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, (1992), pp. 18-34.
- [7] T. Sikora and B. Makai, "Shape-adaptive DCT for generic coding of video", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 1, (1995), pp. 59-62.
- [8] K. Belloulata, R. Stasinski and J. Konrad, "Region-based image compression using fractals and shape-adaptive DCT", *Proceedings of the IEEE International Conference on Image Processing, Kobe, Japan, vol. 2, (1999) October 24-28*, pp. 815-819.

Authors



Yamin Zheng was born in Xi'an in 1961. She received her B.E. degree from Xi'an Jiaotong University, China, in 1982 and her M.E. degree in circuit, signal and systems from Beihang University, China, in 1989. Now she is an associate professor in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. Her current research interest includes electronic technology, computer technology and its applications.



Xiaoyan Wang was born in Shandong province, China in 1990. She received her B.S. degree in electronic information science and technology from Shandong University, Weihai, China, in 2012. Now she is pursuing her M.E. degree in circuits and systems in Shandong University, Weihai, China. Her research interests concentrate on image and video processing.



Chengyou Wang was born in Shandong province, China in 1979. He received his B.E. degree in electronic information science and technology from Yantai University, China, in 2004 and his M.E. and Ph.D. degree in signal and information processing from Tianjin University, China, in 2007 and 2010 respectively. Now he is a lecturer in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. His current research interests include digital image and video processing, wavelet analysis and its applications, and smart grid technology.