DE GRUYTER OPEN Open Comput. Sci. 2015; 5:13–21

Research Article Open Access

Pooneh Bagheri Zadeh*, Akbar Sheikh Akbari, and Tom Buggy

DCT image codec using variance of sub-regions

DOI 10.1515/comp-2015-0003

Received May 20, 2011; accepted July 21, 2015

Abstract: This paper presents a novel variance of sub-regions and discrete cosine transform based image-coding scheme. The proposed encoder divides the input image into a number of non-overlapping blocks. The coefficients in each block are then transformed into their spatial frequencies using a discrete cosine transform. Coefficients with the same spatial frequency index at different blocks are put together generating a number of matrices, where each matrix contains coefficients of a particular spatial frequency index. The matrix containing DC coefficients is losslessly coded to preserve its visually important information. Matrices containing high frequency coefficients are coded using a variance of sub-regions based encoding algorithm proposed in this paper. Perceptual weights are used to regulate the threshold value required in the coding process of the high frequency matrices. An extension of the system to the progressive image transmission is also developed. The proposed coding scheme, JPEG and JPEG2000 were applied to a number of test images. Results show that the proposed coding scheme outperforms JPEG and JPEG2000 subjectively and objectively at low compression ratios. Results also indicate that the proposed codec decoded images exhibit superior subjective quality at high compression ratios compared to that of JPEG, while offering satisfactory results to that of JPEG2000.

Keywords: discrete cosine transform; image compression; perceptual weights; quad-tree coding; variance of sub-regions

***Corresponding Author: Pooneh Bagheri Zadeh:** Bagheri Zadeh School of Computer Science and Informatics, De Montfort University, U.K., Email: pooneh.bagherizadeh@dmu.ac.uk
Akbar Sheikh Akbari: School of Computing, Creative Technology and Engineering, Faculty of Arts, Environment and Technology, Leeds Beckett University, U.K., Email: a.sheikh-akbari@leedsbeckett.ac.uk
Tom Buggy: Division of Communication, Network and Electronic Engineering, School of Engineering & Computing, Glasgow Caledonian University, 70 Cowcaddens Road, Glasgow G4 0BA, UK E-mail: t.buggy@gcal.ac.uk

1 Introduction

With advances in multimedia technologies and internet applications, demand for transmission and storage of voluminous amounts of multimedia data have dramatically increased. Over the past two decades much effort has been invested in developing efficient image compression techniques and has led to the development of many image compression algorithms[1] [1, 2]. In recent years wavelet based image coding schemes have achieved impressive success, mainly due to the novel approaches taken by these schemes in data organization and representation of wavelet-transformed coefficients [1, 2]. Wavelet based methods offer very high compression efficiency in terms of PSNR. However, the high compression efficiency in the rate/distortion sense does not necessarily correspond to a similar degree of compression efficiency in terms of visual quality versus bitrates [3, 4]. Grgic *et al.* [5] has shown that Discrete Cosine Transform (DCT) produces slightly better results than wavelets especially at low compression ratios while its computational complexity is less expensive than that of wavelets. A comparative study on wavelets and DCT, reported by Xiong *et al.* [6], shows that the main factors to distinguish image compression schemes are the way the transformed coefficients are re-arranged, quantized and coded rather than the difference between the transforms used. Although wavelets appear to provide more flexible space-frequency resolution tradeoffs than DCT, the latter has been successfully employed as the first step in several coding algorithms such as: JPEG[1], MEPG[2] and H.26x[3] because of its compression capabilities and processing speeds.

JPEG, the first image compression standard[1], divides an image into blocks of 8x8. The resulting blocks are then decorrelated using a DCT. The transformed coefficients are

1 Information Technology-JPEG-Digital Compression and Coding of Continuous-Cone Still Image-Part 1: Req. and Guidelines, 1994. ISO/IEC 10918-1 and ITU-T Recom. T.81.

2 ISO/IEC JTC1/SC29/WG11. MPEG-4 Overview, available from http://mpeg.telecomitalialab.com/standards/mpeg-4/mpeg-4.htm#E25E3, March 2001.

3 ITU-T Recommendation H.263, Video codec for Low bit rate communication, Version 1, November 1995; Version 2, January 1998, Version 3, November 2000.

zigzag scanned, quantized and huffman coded. JPEG provides good compression ratios, particularly when dealing with natural images. However, the decoded image using JPEG suffers from blocking artifacts especially in smooth regions, and the Gibbs phenomenon, which causes ripples or oscillations around sharp edges at high compression ratios [7]. Recent DCT-based encoders with novel approaches to data organization and representation of DCT coefficients have obtained higher compression efficiency compared to JPEG [4, 8, 9]. A DCT-based embedded image coding (EZDCT) algorithm was proposed by Xiong *et al.* [8], which like JPEG, divides an image into blocks of 8x8 and performs DCT on each block. Each resulting DCT block is viewed as wavelet decomposition with a uniform 8x8 subband decomposition. The DCT coefficients are then rearranged into a 3-level wavelet pyramidal structure and the re-arranged coefficients are coded using embedded zerotree coding scheme. Another DCT-based image codec that uses the morphological representation of DCT coefficients (MRDCT) was presented by Debin *et al.* [4]. It re-arranges DCT coefficients like the EZDCT and then uses a morphological dilation to represent the re-arranged DCT coefficients. An Embedded Quadtree DCT based image compression scheme (EQDCT) was suggested by Xing-Song *et al.* [9]. It re-arranges the DCT coefficients in a similar manner to the EZDCT scheme and then uses an adaptive quadtree splitting algorithm, which is similar to the quad tree coding strategy of EZBC in wavelet domain, to code the significant DCT coefficients. They reported superior performance to that of EZDCT, MRDCT and JPEG at high compression ratios.

Variance of the image data have been used in a number of image compression techniques and offers promising visual quality especially at high compression ratios [10, 11], while the application of variance of the transformed image data in image compression has been less reported in the literature. In progressive image transmission, an image is transmitted in two or more stages. A coarse but recognizable image with basic quality is first transmitted to the receiver. The transmission of further details is made at the later stages upon the request of the viewer. Some image coding schemes, e.g. JPEG, JPEG2000, SPIHT and EZW, are very suitable for progressive image transmission due to the good energy compaction properties of their transforms, i.e. Wavelets and DCT [12].

In this paper, a novel Variance of Sub-regions and DCT (VSDCT) based image-coding scheme is presented. The proposed coding scheme divides the input image into a number of non-overlapping blocks. It applies a DCT on coefficients in each block and transforms the coefficients into their frequency domain. The coefficients with the same frequency index at different DCT blocks are grouped together to make a number of matrices. The matrix containing the DC coefficients is losslessly coded. The matrices containing high frequency coefficients are coded using a novel variance of sub-regions based encoding algorithm. The proposed encoder applies a hierarchical estimation algorithm to code the coefficients in each matrix. The hierarchical estimation algorithm assumes that the distributions of the coefficients in the matrices are Gaussian in some regions. A threshold on the variance of the coefficients is used to determine if it is possible to estimate the coefficients in the input matrix with a single Gaussian distribution or if it needs further dividing into four sub-blocks. This hierarchal algorithm is repeated until the distribution of the coefficients in all sub-blocks fulfils the above criteria. Finally, the mean value of the Gaussian distribution of each block is taken as an estimation value for all coefficients in that block. During the encoding process a quadtree-like binary map is generated to save a record of the hierarchical operation, which is used in decoding process. The application of the codec to progressive image transmission is also investigated. The rest of the paper is organized as follows: in Section 2 the proposed coding scheme is discussed; Section 3 explains the progressive mode of operation. Section 4 presents comparative experimental results; finally Section 5 concludes the paper.

2 Variance of sub-regions and DCT based image encoder

A block diagram of the Variance of Sub-regions and Discrete Cosine Transform (VSDCT) based image encoder is illustrated in Figure 1. A gray scale image is input to the encoder. The encoder divides the input image into a number of 8x8 non-overlapping pixel blocks called B_{11} to B_{nn} as shown in Figure 1(a). The coefficients in each block are then transformed into the frequency domain using a discrete cosine transform as shown in Figure 1 (b) where A_{0-ij} to A_{63-ij} are DCT transformed coefficients in the B_{ij} block. The transformed coefficients with the same frequency index at different blocks are then grouped together generating 64 matrices called M_0 to M_{63}, where M_0 contains the DC coefficients and M_1 to M_{63} contain the AC coefficients from the lowest to the highest frequency, respectively.

Figure 1(c) shows one of these matrices (M_k), where A_{k-11} to A_{k-nn} in this matrix represent coefficients with the frequency index (k), which k can take a value between 1 and 63,at different transformed blocks. In this figure, indices 11 to nn represent the block that the coefficients be-

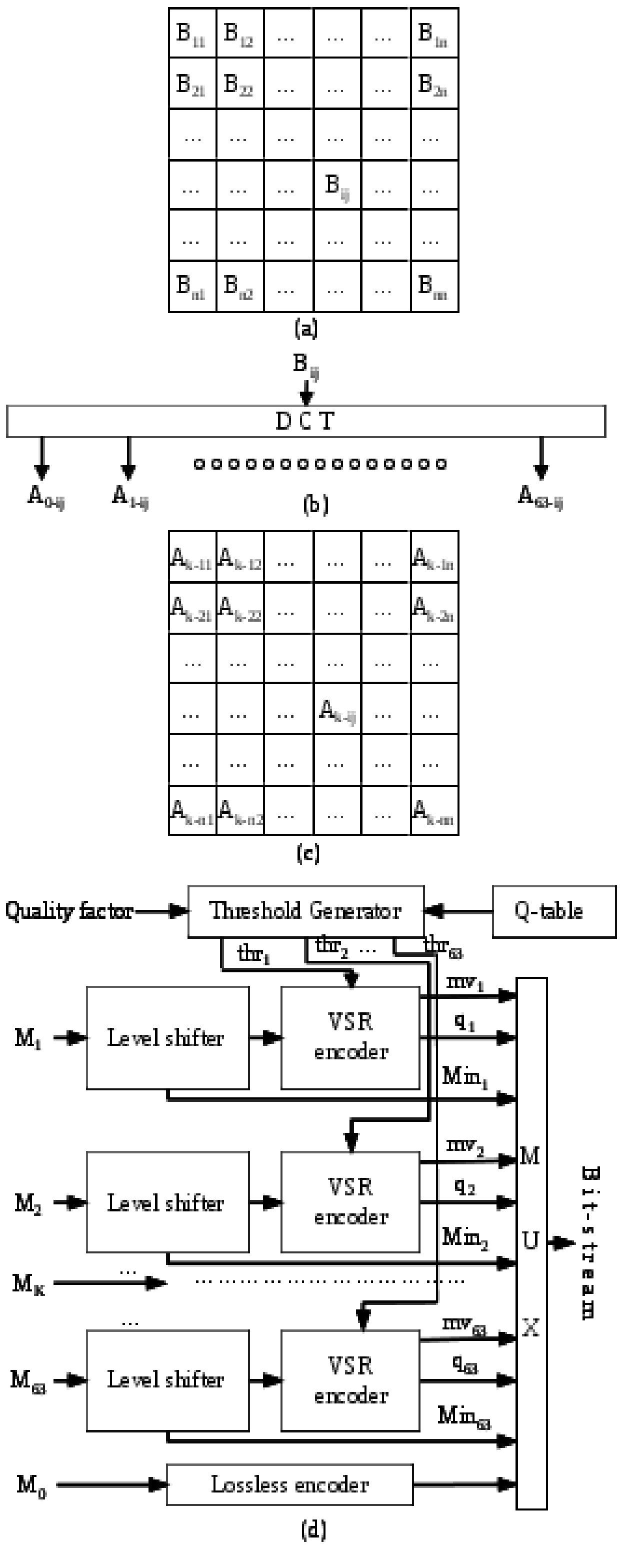

Figure 1: Variance of sub-regions and DCT based image encoder (a) non–overlapping blocks of an image, (b) DCT transformed coefficients of block ij, (c) organization of the coefficients in matrix M_k (d) encoding algorithm, where VSR stands for Variance of Sub-Regions.

long to. Figure 1(d) illustrates the encoding stage of these 64 matrices. The M_0, which contains most of the image energy, is losslessly coded, using lossless Differential Pulse Code Modulation (DPCM) method. The M_1 to M_{63} matrices are independently coded using the following operations: (i) Coefficients in each matrix are first level shifted to have a minimum value of zero; (ii) the resulting coefficients are then coded using a novel Variance of Sub-Regions (VSR) based encoding algorithm, which is presented in Section B. The statistical encoder takes coefficients in each matrix and a threshold value, generated specifically for that matrix (detailed in Section A), and performs the encoding process. The output of this encoder is a mean vector (**mv**), which carries the mean values, and a binary vector (**q**), which carries its quadtree-like data. (ii) Finally a multiplexor puts the encoded information together and generates the output bitstream.

2.1 Threshold generation

The threshold value for each matrix is generated using a uniform quality factor and its JPEG quantization step, which is exploited from the JPEG quantization table. JPEG quantization table takes into account the properties of the human visual system for different spatial frequencies. The threshold values are calculated using the following empirical formula:

$$Threshold = (Q_{step} \times Quality_{factor})^{1/4}, \tag{1}$$

where the Q_{Step} is its related JPEG quantization step [13], and the $Quality_{factor}$ which takes any positive values between 0 and 1000, controls the compression ratio. This empirical formula has been found to be appropriate for a wide range of test images, e.g. Lena, House, Elaine, Bee, Boat, Goldhill, map, peppers, Zelda, Café. An example of the JPEG quantization table can be found in [14].

2.2 Variance of sub-regions based encoder

The block diagram of the new Variance of Sub-Regions (VSR) based encoding algorithm is shown in Figure 2. The variance of sub-regions based encoder takes the coefficients of one of the high frequency matrices (M_1 to M_{63}) and a threshold value, generated specifically for that matrix (detailed in Section 2.1), and performs the encoding process on it. For simplification the input matrix in explanation of the encoder is called I. The encoding process for the input matrix I is as follows:

The encoder first defines two empty vectors called **mv** (mean value vector) and **q** (quadtree-like vector). It then calculates the variance and the mean value of the matrix I and compares the resulting variance value with the threshold value. If the variance is less than the threshold value, the matrix is coded by its mean value m and one bit binary data equal to 0, which are placed in the **mv** and **q** vectors, respectively. Otherwise one bit binary data equal to one is placed at the **q** vector and the size of the matrix is checked. If the size of the matrix is 2x2, the four coefficients of the subband are scanned and placed in the **mv** vector and encoding process is ended by sending the mean value vector **mv** and the quadtree-like vector **q**. If the size of the matrix is greater than 2x2, the matrix I is divided into four equal non-overlapped blocks. These four blocks are then processed from left to right, as shown in the block diagram. For simplification, the continuation of the coding process of the first block, I_1, is discussed. This process is exactly repeated on the three other blocks. Processing of the first block I_1 is described as follows:

The variance and the mean value of the sub-matrix I_1 is first calculated and then the resulting variance value is compared with the input threshold value. If it is less than the threshold value, the calculated mean value m_1 is concatenated to the mean value vector **mv** and one bit binary data equal to 0 is appended to the quadtree-like vector **q**. The encoding process of this sub-block is terminated at this stage. Otherwise, the size of the sub-block is checked. If it is 2x2, one bit binary data equal to 1 is appended to the current quadtree-like vector **q** and the four coefficients of the sub-block are scanned and concatenated to the **mv** vector and encoding process is ended for this sub-block. If its size is greater than 2x2, one bit binary data equal to 1 is concatenated to the current quadtree-like vector **q** and the sub-block I_1 is then divided into four equal non-overlapped blocks. These four new sub-blocks, which are named successor sub-blocks, are processed from left to right in the same way that their four ancestor sub-blocks were encoded.

The above process is continued until whole successor blocks are encoded. When the encoding process is finished two vectors **mv** and the **q** are passed to the output.

3 Progressive Variance of sub-regions and DCT based image encoder

A block diagram of the progressive Variance of Sub-regions and DCT (VSDCT) based image encoder is shown in Figure 3. It can be seen from Figure 3 that the encoding process is similar to that explained for the variance of sub-regions and DCT based image encoder in Section 2, apart from the bitstream generator. It employs several multiplexers to generate a signal to noise scalability, which supports the progressive image transmission. Figure 3d shows that each multiplexer puts the encoded information for each matrix (a mean vector (**mv**), a binary vector (**q**), and a minimum value (Min)) together generating a bitstream for that matrix. This bitstream is called BS_K, where k is the frequency index that specifies the correspondent matrix. The resulting bitstreams are transmitted from BS_0 to BS_{63}, respectively to perform progressive image transmission.

Figure 4 shows a block diagram of the progressive variance of sub-regions and DCT based image decoder. The decoding process is started when the reception of the BS_0, which contains the DC coefficients of all transformed DCT blocks, is completed. The decoder then assumes that the information in the remaining matrices is zero and reconstructs the output image with the basic quality using the received data. The decoding process is then continued as follow: (i) it waits until the reception of data for the next matrix is completed; (ii) it assumes that data in the remaining matrices are zero and reconstructs the image using information in the received matrices; (iii) if the reception of the information for all matrices is not completed, it goes back to stage (i). This process is repeated until an image of the desired quality is obtained.

4 Experimental results

In order to evaluate the performance of the proposed Variance of Sub-regions and DCT (VSDCT) based image coding scheme, three standard 8-bit greyscale images of resolution 512x512, 'Lena', 'House' and 'Elaine' were coded using JPEG, JPEG2000 and the proposed coding scheme. The PSNR measurements for the encoded images using the three techniques at different compression ratios are shown in Figure 5. Results indicate that the proposed coding scheme outperforms JPEG and JPEG2000, objectively at low compression ratios while offers inferior performance at higher compression ratios. However, it is well known that the PSNR is an unreliable metric for measuring the visual quality of the compressed images [15]. Consequently, to illustrate the visual quality obtained using the VSDCT, JPEG and JPEG2000 encoding schemes, a zooming view of 'Lena' image at compression ratios of 5, 15 and 40 are shown in Figures 6. From Figure 6(i), which shows Lena test image at compression ratio of 5, it can be seen that the VSDCT image has slightly higher visual quality than JPEG image, while VSDCT image exhibits almost the same

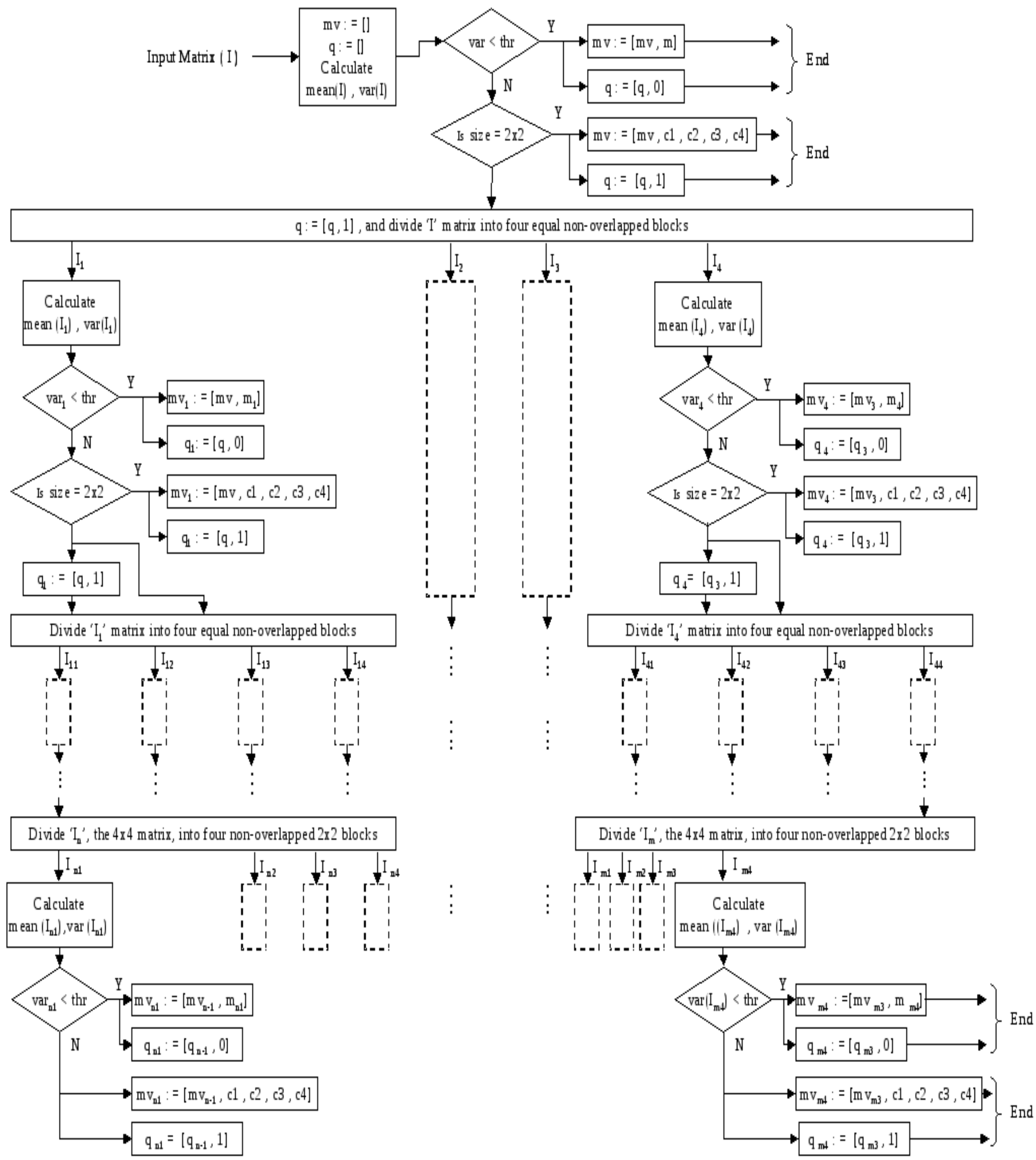

Figure 2: Block diagram of the variance of sub-regions based encoder.

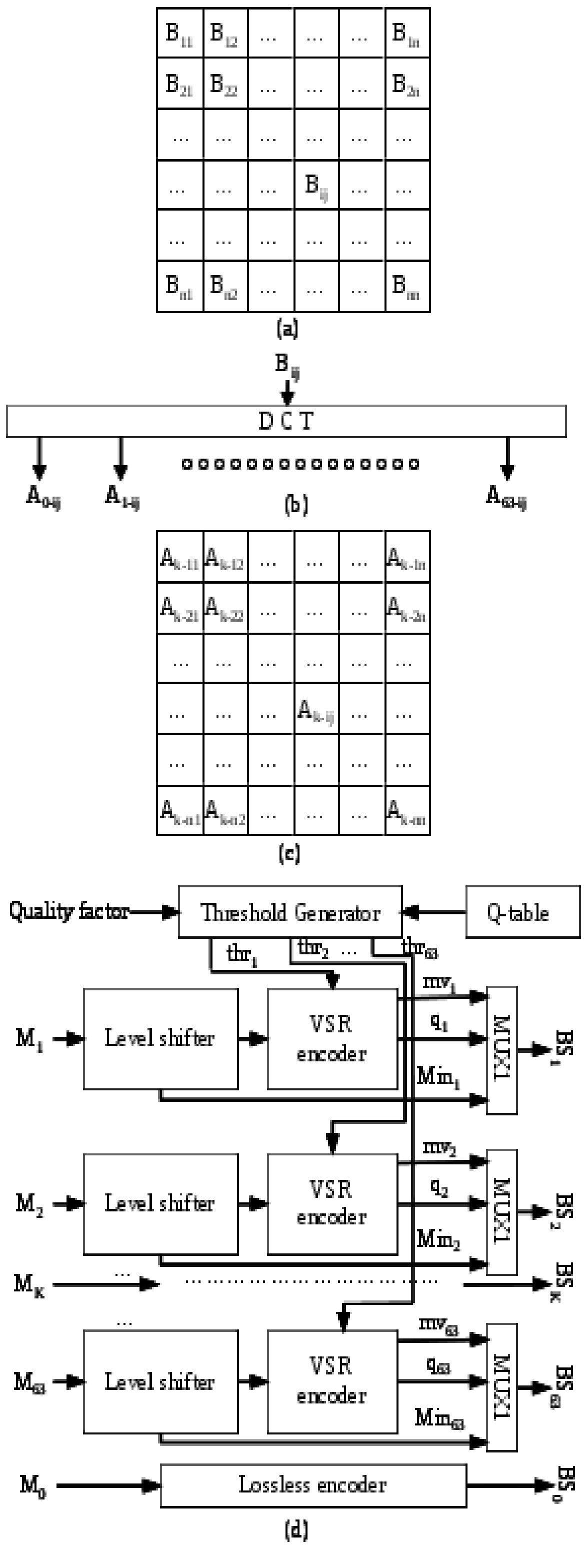

Figure 3: Progressive Variance of Sub-regions and DCT based image encoder: (a) non–overlapping blocks in an image, (b) DCT transformed coefficients of the block ij, (c) organization of the coefficients in matrix M_k (d) encoding algorithm.

visual quality compared to the JPEG2000 image. From Figure 6(ii), which illustrates Lena test image at compression ratio of 15, blocking artefacts on the shoulder and some regions in the background of the JPEG image are noticeable, while blocking artefacts hardly visible in the VSDCT image. The visual quality of the JPEG2000 image at this compression ratio is slightly higher than VSDCT image. From Figure 6(iii), which shows Lena test image at compression ratio of 40, it can be seen that JPEG image suffers from severe blocking artefacts, which limits the application of the JPEG codec in coding images at high compression ratios. From this figure, it can be observed that the VSDCT image exhibits moderate blocking artefacts. In overall, the VSDCT image has superior visual quality to that of JPEG, while it exhibit acceptable visual quality compared to the JPEG2000 image.

In summary, the VSDCT technique offers superior visual quality to that of JPEG at all compression ratios. This superiority does not come without a price. The computational complexity of the proposed codec seems to be higher than that of JPEG. In the proposed codec the DCT coefficients are first grouped into 64 matrices and then each matrix is coded independently using a variance of subregions based encoder while in JPEG the transformed coefficients are first quantized and then entropy coded. It can be seen from Figures 6 that the VSDCT image codec gives comparable visual quality to that of JPEG2000 at compression ratios below 30, while it offers an acceptable visual quality at higher compression ratios. The proposed coding scheme also offers the following advantages: (i) its architecture facilitate its parallel implementation, as matrices are coded independently (iii) its bitstream is efficient for unequal error protection, which offers higher error protection at lower bitrates when transmitting in noisy environments. To evaluate the performance of the codec in the progressive mode of operation, Lena test image was progressively coded using the proposed coding algorithm. Figure 7 illustrates the resulting images at stages 1, 2, 5, 25, 45 and 64 for the proposed codec. The bitrates for the illustrated six quality stages are 0.09, 0.17, 0.4, 0.7, 2.3 and 2.6 bits per pixel (bpp) respectively. Although, at the earlier stages images contain visible coding distortion, they can conveniently be used for browsing and evaluation.

5 Conclusion

In this paper a new variance of sub-regions and DCT based image coding scheme was presented. It divides the input image into a number of non-overlapping pixel blocks. The

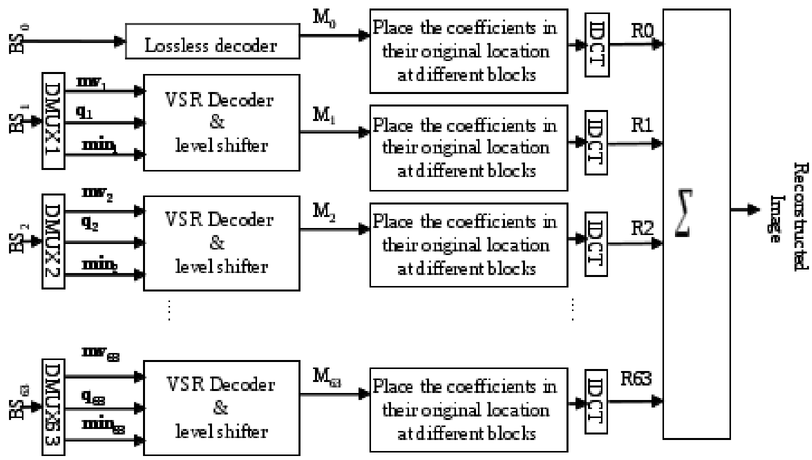

Figure 4: Block diagram of the progressive variance of sub-regions and DCT based image decoder, where VSR stands for Variance of Sub-Regions.

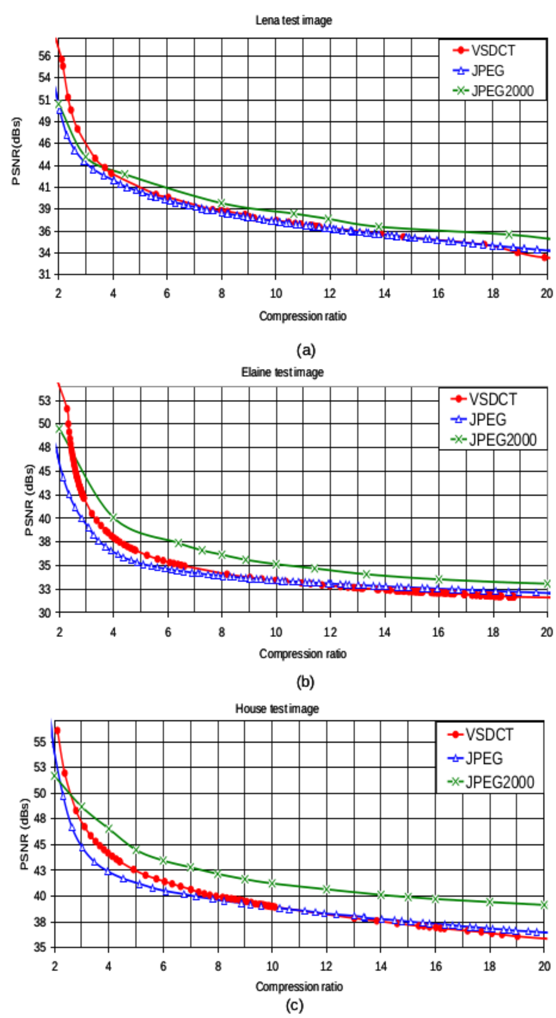

Figure 5: Comparison of coding performance for the proposed VSDCT, JPEG and JPEG2000 image compression schemes: (a) Lena, (b) House, and (c) Elaine.

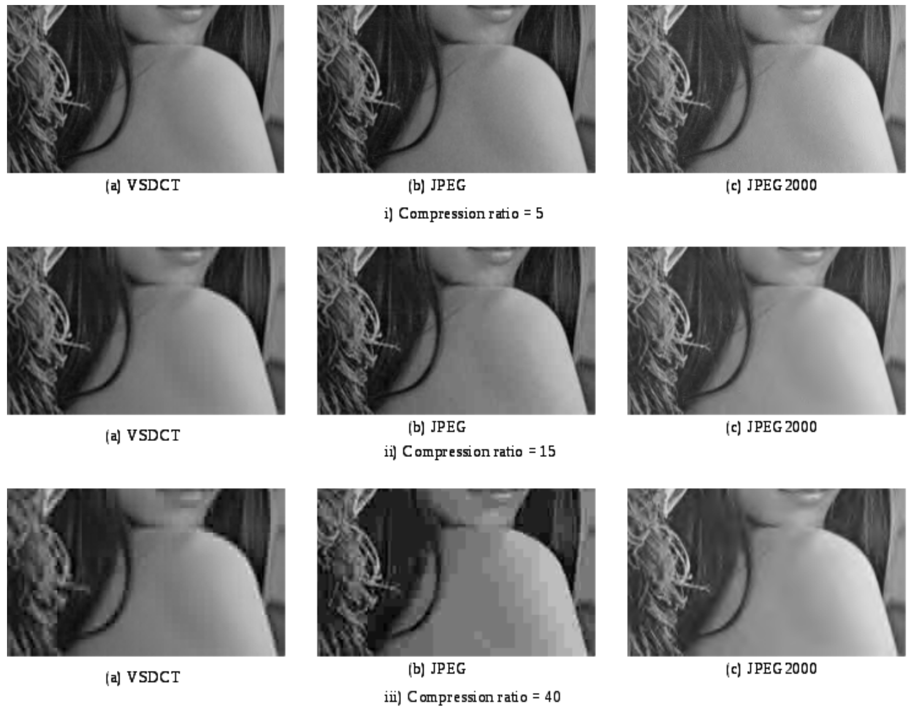

Figure 6: Zooming view of Lena test image at compression ratio of, i) 5, ii) 15 and iii) 40 encoded by (a) VSDCT image codec (b) JPEG, and (c) JPEG2000.

Figure 7: Coding Reconstructed "Lena" test image at the stages (a) 1, (b) 2, (c) 5, (d) 25, (e) 45 and (f) 64.

resulting blocks are then decorrelated using a DCT transform. The coefficients with the same frequency index from the transformed blocks are put together generating a number of matrices. The matrix contains DC coefficients was losslessly coded. The remaining matrices were coded using a novel Variance of Sub-Regions (VSR) based encoder. The VSR encoder estimates the input matrices with the mean value of a number of symmetrical 2D Gaussian waveforms. Results have shown that the proposed codec outperforms JPEG subjectively at all compression ratios. The results have also indicated that the VSDCT offer comparable subjective quality to JPEG2000 at low compression ratios.

References

[1] J.M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE T. Acoust. Speech. 41(12), 3445–3462, 1993
[2] A. Said, W.A. Pearlman, A new, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees, IEEE Trans. Circuits Syst. Video Technol. 6(3), 243–250, 1996.
[3] Z. Xiong, K. Ramchandran, M. Orchard, Space-frequency quantization for wavelet image coding, IEEE Trans. Image Process. 6(5), 677–693, 1997
[4] Z. Debin, W. Gao, Y.K. Chan Morphological representation of DCT coefficients for image compression IEEE Trans. Circuits Syst. Video Technol. 2(9), 819–823, 2002
[5] S. Grgic, M. Grgic, B. Zovko-Cihlar, Performance Analysis of Image Compression using Waveletsl, IEEE Trans. Ind. Electron., 48(3), 2001
[6] Z. Xiong, K. Ramchandran, M.T. Orchard, Y. Zhang, A Comparative Study of DCT- and Wavelet-Based Image Coding, IEEE Trans. Circuits Syst. Video Technol. 9(5), 692–695, 1999
[7] F. Alter, S.Y. Durand, J. Froment, Deblocking DCT-based compressed images with weighted total variation, ICASSP 2004, May 2004
[8] Z. Xiong, O. Gulelyuz, M. Orchard, A DCT-based embedded image coder, IEEE Signal Process. Lett. 3(11), 289–290, 1996
[9] H. XingSong, L. Guizhong, Z. YiYang, Embedded quadtree-based image compression in DCT domain, IEEE International Conference on Acoustics, Speech, and Signal Processing 2003 (ICASSP '03) 6, 277–280, 2003
[10] Z.M. Lu, B. Yang, S.H. Sun, Image Compression Algorithms Based On Side-Match Vector Quantization With Gradient-Based Classifiers, IEICE T. Inf. Syst. E85-D(9), 1409–1415, 2002
[11] S. Saryazdi, M. Jafari, A High Performance Image Coding Using Uniform Morphological Sampling, Residues Classifying, and Vector Quantization, EurAsia-ICT 2002, Shiraz, Iran, October 2002
[12] A. Skodras, C. Christopoulos, T. Ebrahimi, The JPEG2000 still image compression standard, IEEE Signal Proc. Mag. 18(5), 36–58, 2001
[13] M. Ghanbari, Video coding an introduction to standard codecs, The Institution of Electrical Engineering, London, UK, 1999
[14] W.B. Pennebaker, J.L. Mitchell, JPEG: Still image compression standard, Van Nostrand Reinhold, New York, 1993
[15] X. Kaia, Y. Jiea, Z.Y. Minb, L.X. Lianga, HVS-based medical image compression, Eur. J. Radiol. 55, 139–145, 2005