Context-Based Rate Distortion Estimation and Its Application to Wavelet Image Coding

HSI-CHIN HSIN [1] TZE-YUN SUNG [2]

[1] Department of Computer Science and Information Engineering
National United University
1, Lien-Da, Miao-Li, 36003, Taiwan
hsin@nuu.edu.tw
[2] Department of Microelectronics Engineering
Chung Hua University
707 Wu-Fu, Hsinchu 30012, Taiwan
bobsung@chu.edu.tw

Abstract: - Embedded image coding in wavelet domain has drawn a lot of attention. Among noteworthy algorithms is the embedded block coding with optimized truncation (EBCOT) algorithm, which has been adopted by the JPEG2000 standard. EBCOT is a two-tier algorithm. Tier-1 is composed of bit-plane coding followed by entropy coding. Tier-2 performs the so-called post compression rate distortion optimization (PCRD), which requires a large memory space for storing all the code streams of code blocks; however, some code blocks of less importance might not be needed for the optimal decoded image at a bit rate. To avoid waste of computational power and memory space, a simple context based rate distortion estimation (CBRDE) is thus proposed to arrange the scanning order of code blocks in an adaptive manner. The CBRDE algorithm is based on the MQ table of JPEG2000, which is available at both encoder and decoder. As a result, there is no need to store and transmit the rate distortion information of code blocks. Experimental results show that the rate distortion curves are almost convex; this demonstrates the potential of CBRDE for embedded wavelet image coding.

Key-Words: - embedded image coding; wavelet transform; JPEG2000; EBCOT; PCRD; CBRDE

1 Introduction

For many applications, image compression is still in great demand [1]. The commonly used discrete cosine transform (DCT) based image coder called Joint Photographic Expert Group (JPEG) shows satisfactory compression results at moderate to high bit rates [2]. Wavelet transform, which has been adopted by the JPEG2000 standard [3], provides many desirable properties, e.g. multiresolution analysis, progressive transmission and embedded coding [4]-[6]. In wavelet domain, the higher detail information of signals is projected onto the shorter basis function with higher spatial resolution; the lower detail information is projected onto the larger basis function with higher spectral resolution; this property matches the characteristics of Human Visual System (HVS) [7]. Image coding with progressive transmission is desirable for Internet streaming and database browsing. By embedded coding, the original image can be coded into a single code stream, from which the decoded image at any bit rate can be obtained. In practice, the embedded code stream of an image is organized in decreasing order of information importance.

By wavelet transform, images can be decomposed into subbands with orientation selectivity. If a wavelet coefficient is insignificant with respect to a given threshold, all the spatially related wavelet coefficients in the higher frequency subbands of the same orientation are likely to be insignificant with respect to the same threshold; this property known as the self similarity of wavelet transform was first exploited by Shapiro to develop the embedded zerotree wavelet (EZW) algorithm [8]. Its improved version: set partitioning in hierarchical trees (SPIHT) is one of the most successful wavelet-based image coders and becomes a benchmark [9]. Mukherjee and Mitra developed a vector extension of SPIHT called VSPIHT, in which wavelet coefficients are first grouped into small vectors and then coded by SPIHT [10].

The National Science Council of Taiwan, under Grant 97-2221-E-239-031-MY2 supported this work.

Whereas the above-mentioned EZW, SPIHT and VSPIHT are quadtree algorithms taking advantages of the self similarity property of wavelet transform, there are other coders based on block coding in wavelet domain, which exploit the energy clustering property of wavelet coefficients. The AGP algorithm divides regions of high energy coefficients into small blocks by quadtree partitioning [11]. The SWEET algorithm takes octave-band partitioning to exploit the pyramid structure of wavelet transform [12]. Pearlman et al. developed the set partitioning embedded block (SPECK) algorithm with both quadtree partitioning and octave-band partitioning [13]. Taubman proposed the embedded block coding with optimized truncation (EBCOT) algorithm [14], which has been adopted by the JPEG2000 standard. In many cases, JPEG2000 is preferable to JPEG at the cost of increasing complexity [15]. For real-time applications, dedicated hardware is often required to implement JPEG2000.

EBCOT is a two-tier algorithm: tier-1 performs embedded block coding with arithmetic coding on each code block of the transform coefficients; tier-2 takes charge of rate distortion control by the post compression rate distortion (PCRD) optimization algorithm. One of the crucial implementation issues of EBCOT is the design of PCRD, which requires a large amount of memory space to store all the code streams of code blocks with their respective rate distortion information. In addition, all the code blocks of an image must be processed before PCRD. However, some code blocks may not be needed to reconstruct the optimal decoded image at a bit rate; this surely leads to waste of computational power. In [16], Fang et al. proposed a precompression rate distortion optimization algorithm to avoid unnecessary computations and reduce the memory space by ignoring the unused code streams.

In this paper, we propose a context based rate distortion estimation (CBRDE) to arrange the scanning order of code blocks in an adaptive manner, such that the available coding bits can be allocated for the most significant code block. The proposed CBRDE is based on the MQ table of JPEG2000, which is available at both encoder and decoder. As a result, the scanning order of code blocks can be also obtained at the decoder, and therefore transmission of the contributions of code blocks can be omitted. The remainder of this paper proceeds as follows. In Section 2, wavelet transform and the EBCOT algorithm are reviewed briefly. Section 3 presents the proposed CBRDE algorithm to arrange the code blocks of images. Experimental results are given in Section 4, and conclusion can be found in Section 5.

2 Wavelet Transform and EBCOT

Wavelet transform provides many desirable properties such as multiresolution analysis, subband decomposition with orientation selectivity, joint space-spatial frequency localization, self similarity across subbands of the same orientation, and energy clustering within each subband. In this section, wavelet transform and the EBCOT algorithm are reviewed briefly.

2.1 Wavelet transform

The wavelet transform of a signal, $S_\ell(n)$, at resolution ℓ is as follows.

$$S_{\ell+1}(n) = \sum_k S_\ell(k)\ h(2n-k)$$
$$D_{\ell+1}(n) = \sum_k S_\ell(k)\ g(2n-k) \tag{1}$$

where $S_{\ell+1}(n)$ is the approximation at the next coarser resolution $\ell+1$, $D_{\ell+1}(n)$ is the detail between resolutions ℓ and $\ell+1$, $h(n) = <\phi, \phi_{-1,-n}>$, $g(n) = <\psi, \phi_{-1,-n}>$, $<\cdot,\cdot>$ is an inner product operator, ψ is a valid (mother) wavelet, ϕ is the corresponding scaling function and $\phi_{-1,-n}(x) = 2^{-1/2}\phi(2^{-1}x-n)$. $S_\ell(n)$ can be exactly reconstructed from $S_{\ell+1}(n)$ and $D_{\ell+1}(n)$ by using the following inverse wavelet transform.

$$S_\ell(n) = \sum_k S_{\ell+1}(k)\ \tilde{h}(n-2k) + \sum_k D_{\ell+1}(k)\tilde{g}(n-2k) \tag{2}$$

where $\tilde{h}(n) = h(-n)$ and $\tilde{g}(n) = g(-n)$.

For image applications, 2-D wavelet transform can be obtained by the tensor product of two 1-D wavelet transforms performed horizontally followed by vertically, or vice versa. Fig. 1 shows an example of 3-level 2-D wavelet transform, where subbands of wavelet coefficients: $D_\ell^1(m,n)$, $D_\ell^2(m,n)$ and $D_\ell^3(m,n)$ denoted by HL_ℓ, LH_ℓ and HH_ℓ represent the detail information at resolution ℓ in the horizontal, vertical and diagonal orientations, respectively; Subband of scaling coefficients $S_3(m,n)$ denoted by LL_3 represents the approximation at the coarsest resolution 3. The original image is usually taken as $S_0(m,n)$ at the finest resolution 0. After 1-level 2-D wavelet transform, $S_\ell(m,n)$ is decomposed into $S_{\ell+1}(m,n)$,

$D^1_{\ell+1}(m,n)$, $D^2_{\ell+1}(m,n)$ and $D^3_{\ell+1}(m,n)$. Similarly, 2-D inverse wavelet transform can be obtained by the tensor product of two 1-D inverse wavelet transforms performed vertically followed by horizontally, or vice versa.

2.2 The EBCOT algorithm

In many cases, the DWT-based JPEG2000 standard outperforms the DCT-based JPEG standard. The idea behind the heart of JPEG2000 known as the EBCOT algorithm is to exploit energy clustering of wavelet coefficients. EBCOT is a two-tier algorithm; tier-1 performs bit-plane coding (BPC) followed by arithmetic coding (AC); tier-2 aims for post compression rate distortion optimization. The quantized wavelet coefficients of an image are coded by a context-based arithmetic coder known as the MQ coder, in which the probability models are stored in the MQ table. Figure 2 depicts block diagram of the JPEG2000 encoder.

In EBCOT, each BPC is further divided into three coding passes, namely significance propagation pass, magnitude refinement pass, and cleanup pass. Four primitive coding operations, namely significance coding operation, sign coding operation, magnitude refinement coding operation, and cleanup coding operation are performed in the corresponding coding passes. For a transform coefficient that is currently insignificant, if any of the 8 neighboring transform coefficients are already significant, it is coded in the significance propagation pass using the significance coding operation; otherwise, it is coded in the cleanup pass using the cleanup coding operation; if this coefficient becomes significant, the sign is coded immediately using the sign coding operation. In the magnitude refinement pass, the magnitude of the significant transform coefficients that have been found in previous coding passes is updated using the magnitude refinement coding operation. The output bit streams of coding passes can be further coded by using an arithmetic coding engine to improve the compression performances at the cost of computational complexity. Based on the current status of the 8 neighboring transform coefficients, JPEG2000 defines 18 context labels, and stores their respective probability modes in the MQ table. More precisely, 10 context labels are defined for the significance coding operation and the cleanup coding operation, 5 context labels for the sign coding operation, and 3 context labels for the magnitude refinement coding operation.

3 The CBRDE Algorithm

In JPEG2000, a large image is first divided into rectangular sub-images called tiles; each tile is decomposed into subbands by wavelet transform; every subband is partitioned further into small blocks called code blocks, which are quantized to form bit-planes and then coded by EBCOT from the most significant bit-plane to the least significant bit-plane. For each bit-plane, all the code blocks of an image must be processed in the first tier of EBCOT before proceeding with the application of the PCRD algorithm in the second tier. As code blocks of less importance are not needed for the optimal decoded image at a given bit rate, waste of computational power and memory space might result. In this section, an efficient scheme is proposed for the scanning order of code blocks such that waste of computational power and memory space can be reduced.

3.1 Adaptive scanning order

Recall that the code blocks of an image are coded bit-plane by bit-plane, from most to least significant, and the output bitstream can be truncated at an intermediate point between bit-planes; it raises the following interesting questions regarding the scanning order of code blocks. For each bit-plane, is there an adaptive scanning order such that the available coding bits can be allocated for the most significant code block? Is there a common piece of information available at both encoder and decoder, based on which code blocks can be arranged adaptively in decreasing order of significance? If so, there is no need to store all the code streams of code blocks and transmit the scanning order from encoder to decoder. Is there any simple way to solve the above-mentioned questions?

As one might expect, the first question can be solved effectively by using the rate distortion curve of code blocks. The optimal scanning strategy turns out to take the code block with the steepest rate distortion slope as the first to be coded. In other words, the available coding bits should be first allocated for the code block with the largest amount of distortion decrease per coding bit. Take the second question into consideration; estimated rate distortion slope is preferable to the true one for arranging the scanning order of code blocks. Since the MQ table of probabilities of more probable symbol (MPS) and less probable symbol (LPS) is available at both encoder and decoder, an efficient, context based rate distortion estimation (CBRDE) is therefore proposed to estimate the rate distortion slope of code blocks, which is given in the following subsection.

3.2 Proposed CBRDE

For each bit-plane, code blocks with significant rates of distortion decrease per coding bit should be coded as early as possible. In order to avoid transmitting the scanning order of code blocks from encoder to decoder, the estimated rate distortion slope of a code block is thus adopted, which is given by

$$S = \frac{E[\Delta D]}{E[\Delta R]} \tag{3}$$

where ΔD and ΔR denote the amount of distortion decrease and the number of coding bits, respectively, and $E[\cdot]$ is an expectation operator. Though most of correlation between images pixels can be removed by wavelet transform, there may still be residual correlation between neighboring wavelet coefficients. To incorporate with the framework of EBCOT, the proposed CBRDE for code blocks is based on the MQ table, which is available at both encoder and decoder. Specifically, let B_i denote the i^{th} code block (in wavelet domain), $b_{ij}(m,n)$ the j^{th} significant bit of a coefficient at position: (m,n) in B_i, and $B_{ij} = \bigcup_{m,n} \{b_{ij}(m,n)\}$. The proposed CBRDE for B_{ij} is as follows.

$$S_{ij} = \frac{\sum_{m,n} prob(b_{ij}(m,n) = 1)}{\sum_{m,n} H(b_{ij}(m,n))} \tag{4}$$

where $H(b_{ij}(m,n))$ denotes the entropy of $b_{ij}(m,n)$ given by

$$\begin{aligned} H(b) = &-prob(b=1)\cdot\log_2(prob(b=1)) \\ &- prob(b=0)\cdot\log_2(prob(b=0)) \end{aligned} \tag{5}$$

The numerator of equation (4) is proportional to the amount of distortion decrease; and the denominator is proportional to the number of coding bits. It is noted that probability models of $b_{ij}(m,n)$ can be obtained directively from the MQ table, and therefore the computational complexity is nothing but one lookup table operation.

Take the commonly used Lena image as an example, the true distortion: ΔD and coding bits: ΔR, and their respective estimates $E[\Delta D]$ and $E[\Delta R]$ (for a given bit-plane) are shown in Fig. 3. As one can see, the nearly proportional relationship between the horizontal axis (true values) and the vertical axis (estimated values) demonstrates the potential of the proposed CBRDE.

3.3 Embedded block coding with CBRDE

Figure 4 depicts the proposed image encoder using embedded block coding (EBC) with adaptive block ordering (ABO). In which, the EBC algorithm is the same as the first tier of EBCOT; however, ABO with CBRDE replaces the second tier of EBCOT, i.e. PCRD. With the adaptation of the MQ table, the code blocks of an image can be arranged according to their respective estimated rate distortion slopes and then coded in an adaptive manner. Figure 5 depicts the image decoder. Recall that the rate distortion slopes of code blocks can be also estimated at decoder. Thus, the overhead of transmitting the scanning order of code blocks from encoder to decoder is unnecessary; this is beneficial to the communication applications.

4 Experimental Results

The proposed CBRDE algorithm for arranging the scanning order of code blocks adaptively is valuated on grayscale images. The test images, namely Lena, Barbara and Fingerprint are shown in Fig. 6. Fingerprint is of size 256×256; Lena and Barbara are of size 512×512. The biorthogonal Daubechies wavelet with 9/7-tap filters is used. The number of decomposition levels is 4. The compression rate is measured in bits per pixel (bpp). The distortion is defined as mean squared error (MSE), which is given by

$$MSE = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (I(m,n) - O(m,n))^2}{M \cdot N} \tag{6}$$

where I and O denote the original image and the decoded image, respectively. The compression rates and MSE values are used to plot the rate distortion curves. To avoid the overhead of transmitting the contributions of code blocks, the second tier of EBCOT, i.e. PCRD is replaced by the CBRDE-based ABO. For comparison, as most of images' energies are concentrated in the low frequency subbands, the first tier of EBCOT incorporated with a fixed scanning order starting from the lowest frequency subband is evaluated. More specifically, the wavelet subbands of an image are coded in a zigzag order from the lowest frequency subband to the highest frequency subband; the code blocks within each subband are scanned from left to right, top to bottom. In addition, for the most significant bit-plane, the initial scanning order of the proposed image coder is the same as the above-mentioned order; however, for the successively less significant bit-planes, the scanning order is adapted bit-plane by bit-plane.

Figure 7 shows the rate distortion curves of Lena image, where the horizontal and vertical axes are the bit rates and MSE values, respectively. As Lena is representative of typical images with a large portion of homogeneous regions, the fixed scanning order from low to high frequency subbands is suitable for coding this image. The rate distortion curve obtained by using the proposed method is almost convex although the improvement is only marginal. For texture-rich Barbara image with a large portion of regular textures, there are many significant wavelet coefficients in the middle-high frequency subbands. As shown in Fig. 8, the coding performance can be improved significantly at low bit rates. Figure 9 shows the simulation results of Fingerprint image. It is noted that the rate distortion curve obtained by the proposed image coder is almost convex; this demonstrates the benefit of the proposed CBRDE when the embedded code stream of an image is truncated at an intermediate point between bit-planes.

5 Conclusion

Wavelet transform has been adopted by JPEG2000 as the underlying method to decompose an image into subbands with orientation selectivity. It provides many desirable properties, e.g. multiresolution analysis, high correlation across wavelet subbands of the same orientation, and energy clustering within each subband; these properties are suitable for the image compression applications. In EBCOT, all the code blocks of an image are coded to generate a set of code streams with their respective contributions to the decoded image, based on which the optimal code stream can be obtained by concatenating the suitably truncated code streams through the PCRD algorithm. As some code blocks, which are less important, are not needed for the optimal decoded image at a given bit rate, waste of computational power and memory space may result. Furthermore, implementation of the PCRD algorithm is one of the crucial issues. To overcome the above-mentioned problems, a simple context-based rate distortion estimation (CBRDE) is proposed to arrange the scanning order of code blocks in an adaptive manner. To avoid transmitting the side information regarding the scanning order of code blocks from encoder to decoder, the proposed CBRDE is based on the MQ table, which is available at both encoder and decoder. The second tier of EBCOT, i.e. PCRD can therefore be replaced by the CBRDE-based adaptive block ordering (ABO). Experimental results show that the rate distortion curves are almost convex.

Recall that for each code block, the coding procedure proceeds in distinct passes. Thus, the proposed image coding with adaptive block ordering can be extended to image coding with adaptive pass ordering, however, at the cost of increasing complexity from the implementation point of view.

References:

[1] K. R. Rao and J. J. Hwang, Techniques and Standards for Image, Video and Audio Coding, Prentice Hall, 1996.

[2] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Data Compression Standards, New York: Van Nostrand, 1993.

[3] JPEG 2000 Image Coding System, ISO/IEC CD15444-1: 1999 (Version 1.0).

[4] M. H. Perng and H. H. Lin, "Image Compression in The Wavelet Domain Using An AR Texture Model with Compressed Initial Conditions," WSEAS Trans. On Signal Processing, vol. 2, pp. 161-166, 2006.

[5] X. Wang, H. Lu and D. Huang, "Research and Application for the Image Compression Algorithm based on Multilayer Clustering Wavelet," WSEAS Trans. On Computers, pp. 1936-1941, 2006.

[6] Q. Li, Z. M. Wang, W. M. Zuo and J. You, "An Efficient Image Compression Scheme Based on PCA and Wavelet Analysis with Application on Distance Learning Data," WSEAS Trans. On Signal Processing, vol. 2, pp. 422-428, 2006.

[7] G. Strang and T. Nguyen, Wavelets and Filter Banks. Wellesley, MA: Wellesley-Cambridge, 1996.

[8] J. M. Shapiro, "Embedded Image Coding Using Zero-Trees of Wavelet Coefficients," IEEE Trans. On Signal Processing, vol. 40, pp. 3445-3462, 1993.

[9] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans. On Circuits Syst. Video Tech. vol. 6, pp. 243-250, 1996.

[10] D. Mukherjee and S. K. Mitra, "Vector SPIHT for Embedded Wavelet Video and Image Coding," IEEE Trans. On Circuits Syst. Video Tech. vol. 13, pp. 231-246, March, 2003.

[11] A. Said and W. A. Pearlman, "Low Complexity Waveform Coding via Alphabet and Sample-Set Partitioning," Proc. SPIE Visual Communications and Image Processing, pp. 25-37, Feb., 1997.

[12] J. Andrew, "A Simple and Efficient Hierarchical Image Coder," Proc. IEEE Int. Conf. Image

Processing (ICIP), vol. 3, pp. 658-661, Oct., 1997.
[13] W. A. Pearlman, A. Islam, N. Nagaraj, and A. Said, "Efficient, Low Complexity Image Coding With a Set-Partitioning Embedded Block Coder," IEEE Trans. On Circuits Syst. Video Tech. vol. 14, pp. 1219-1235, Nov., 2004.
[14] D. Taubman, "High Performance Scalable Image Compression with EBCOT," IEEE Trans. On Image Processing, vol. 9, pp. 1158-1170, July, 2000.
[15] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," IEEE Signal Process. Mag., vol. 18, pp. 36-58, September, 2001.
[16] H.-C. Fang, Y.-W. Chang, T.-C. Wang, C.-T. Huang, and L.-G. Chen, "High-Performance JPEG 2000 Encoder with Rate-Distortion Optimization," IEEE Trans. On Multimedia, vol. 8, no. 4, pp. 645-653, August. 2006.

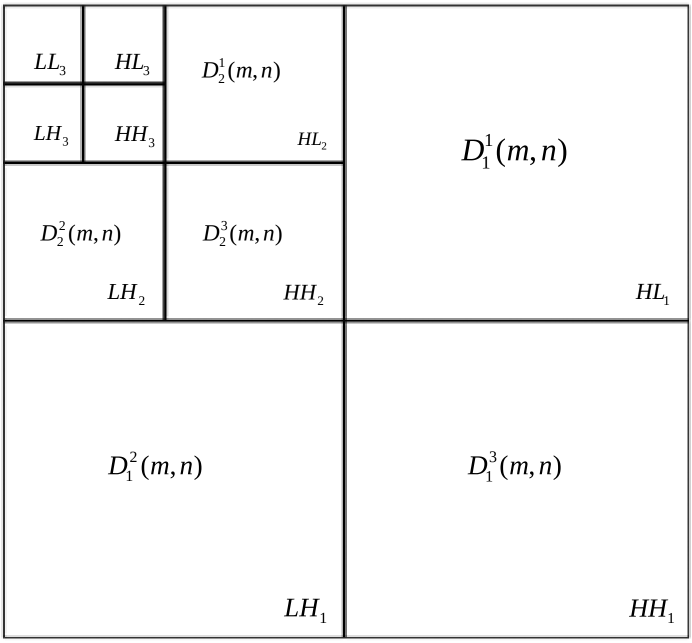


Fig. 1 Example of 3-level 2-D wavelet transform.

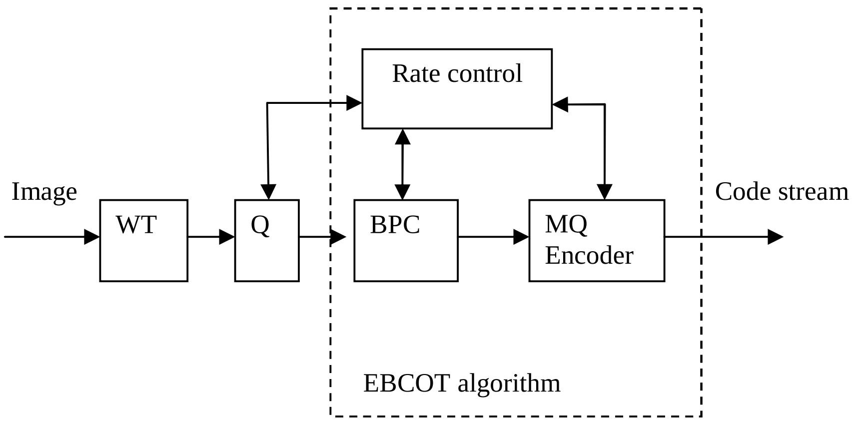
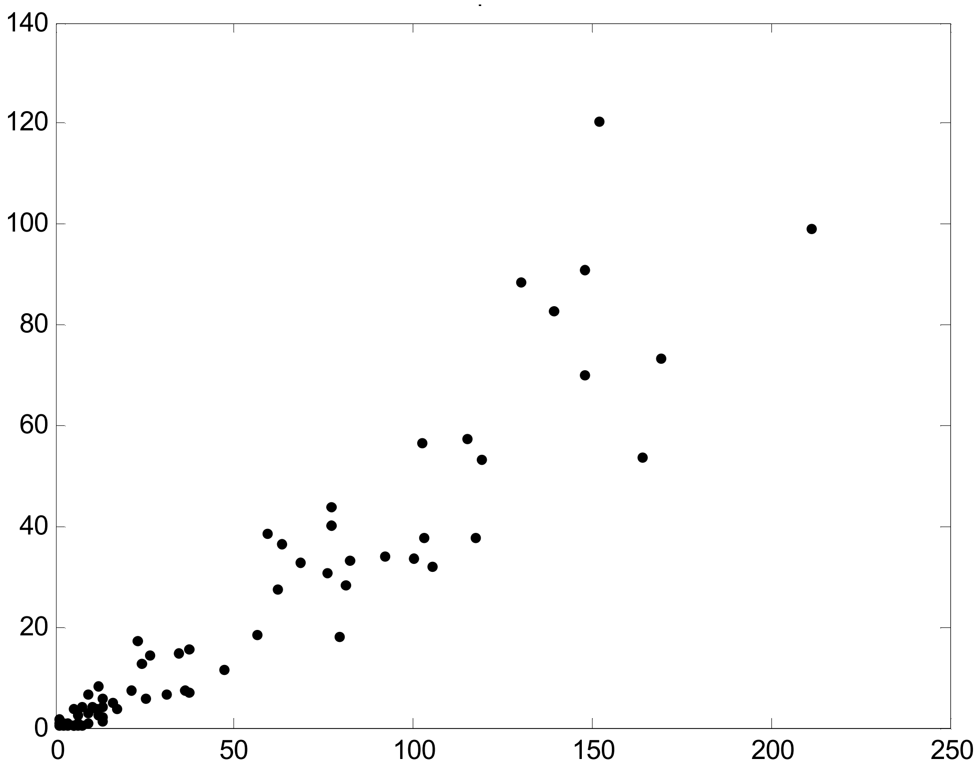
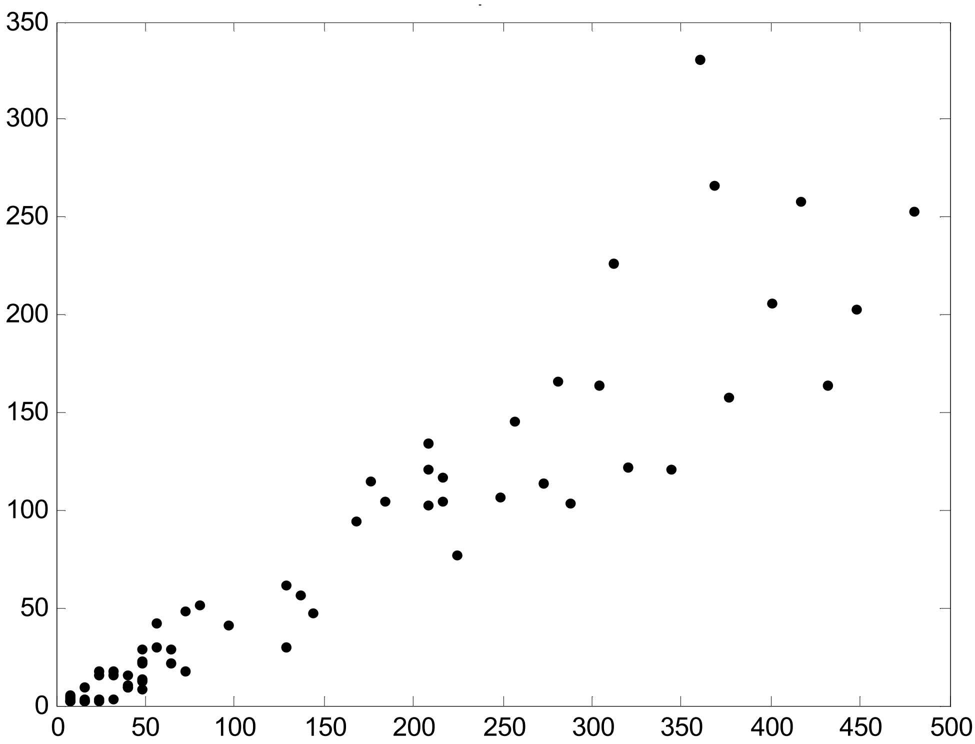


Fig. 2 Block diagram of JPEG2000 encoder



(a)



(b)

Fig. 3 Performance of the proposed CBRDE algorithm applied to Lena image; (a) horizontal axis: true ΔD, vertical axis: estimated $E[\Delta D]$; (b) horizontal axis: true ΔR, vertical axis: estimated $E[\Delta R]$.

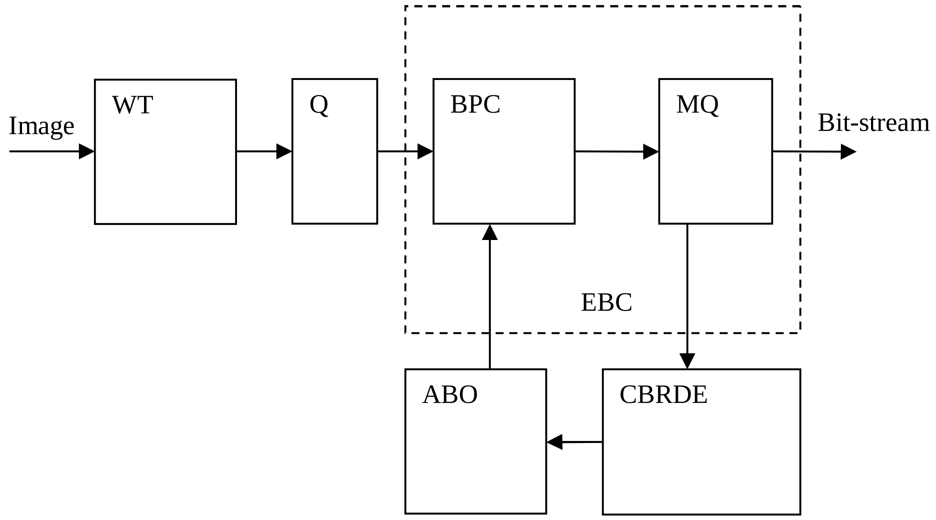


Fig. 4 Proposed image encoder using EBC with ABO

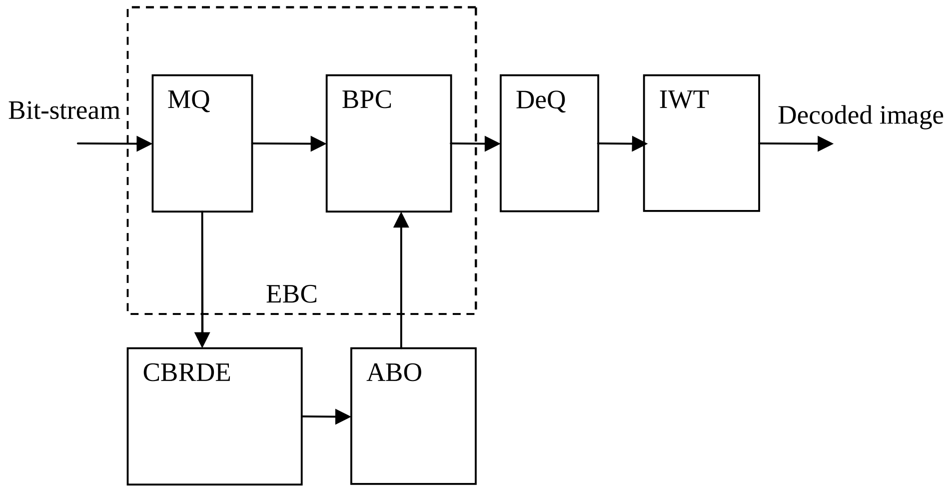


Fig. 5 Proposed image decoder using EBC with ABO

(a)

(b)

(c)

Fig. 6 Test images (a) Lena (b) Barbara (c) Fingerprint

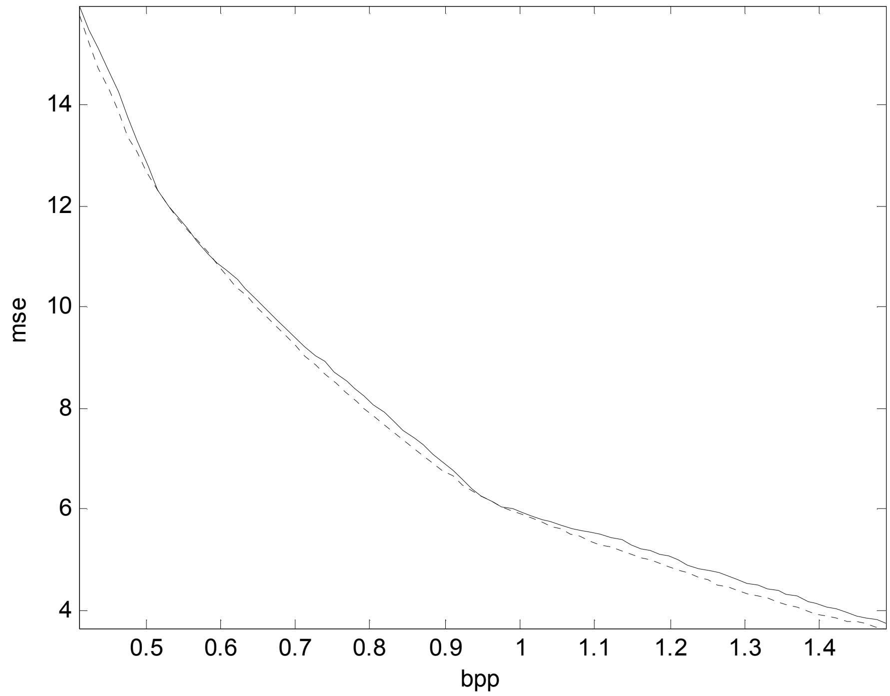


Fig. 7 Rate distortion curves of Lena image by EBC with the CBRDE-based ABO (dashed line) and EBC with a fixed scan order (solid line); Vertical axis: mean square error (MSE); Horizontal axis: bit rate (bpp).

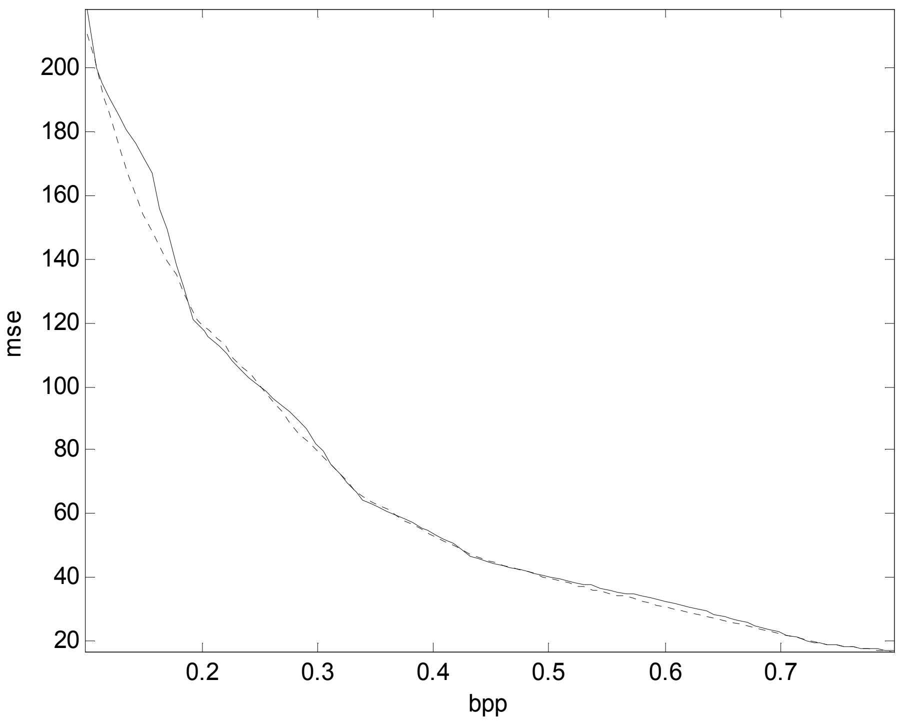


Fig. 8 Rate distortion curves of Barbara image by EBC with the CBRDE-based ABO (dashed line) and EBC with a fixed scan order (solid line); Vertical axis: mean square error (MSE); Horizontal axis: bit rate (bpp).

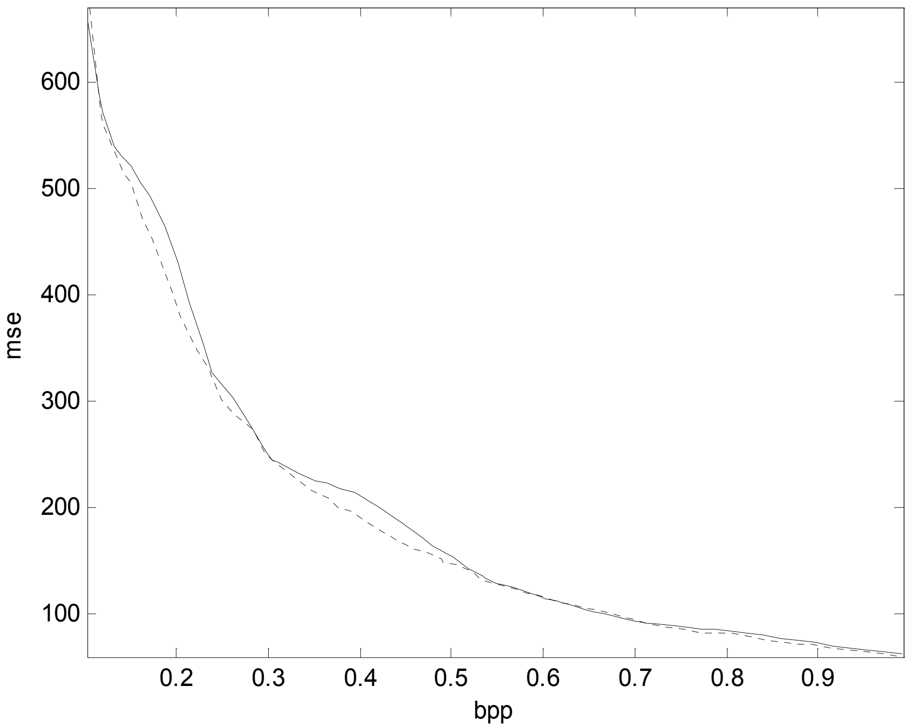


Fig. 9 Rate distortion curves of Fingerprint image by EBC with the CBRDE-based ABO (dashed line) and EBC with a fixed scan order (solid line); Vertical axis: mean square error (MSE); Horizontal axis: bit rate (bpp).