An Edge-Preserving Subband Image Coding Scheme Based on Separate Coding of Region and Residue Sources

Ho-Cheon WEY†, Student Member and Masayuki KAWAMATA†, Regular Member

SUMMARY This paper presents a novel image coding scheme based on separate coding of region and residue sources. In a subband image coding scheme, quantization errors in each subimage spread over the reconstructed image and result in a blurring or a boundary artifact. To obtain high compression ratio without considerable degradation, an input image, in our scheme, is separated into region and residue sources which are coded using different coding schemes. The region source is coded by adaptive arithmetic coder. The residue source is coded using multiresolution subimages generated by applying a subband filter. Each block in the subimages is predicted by an affine transformation of blocks in lower resolution subimages. Experimental results show that a high coding efficiency is achieved using the proposed scheme, especially in terms of the subjective visual quality and PSNR at low bit-rate compression.

key words: subband image coding, separate coding, region source, residue source, multiresolution subimages

1. Introduction

Subband coding (SBC) of still images has received considerable attention from the image coding community for ages [1]–[4]. The multiresolution signal representation given by the SBC scheme provides a convenient framework for both exploiting the specific types of statistical dependencies found in the image and designing quantization strategies. Image coding based on SBC has demonstrated good improvements in coding efficiency over the standard transform-based coder such as the Joint Photographers Expert Group (JPEG) coder.

In the SBC scheme, an input image is divided into a number of subimages, called subbands, using analysis filterbanks. Each subimage is then down-sampled, encoded, and transmitted through a channel. Image compression is obtained as a result of the energy compaction produced by the decomposition of the input image. Generally, high frequency subimages contain less energy and information than low frequency subimages, and therefore they are quantized more coarsely and coded with fewer bits.

However, at low bit-rate compression, the SBC scheme suffers from distortions such as boundary and blurring artifacts. In the SBC scheme, quantization errors of the coefficients in each subimage will spread over the reconstructed image given by the synthesis filterbanks, and then quantization errors appear as a blurring or a boundary artifact. Especially, the boundary artifact is visible around high-contrast contours and can be very annoying for an observer.

To resolve this problem, several schemes have been developed using various choices of analysis/synthesis filters, decomposition structure, and encoding/decoding schemes [4]–[8]. For example, fractal-based SBC schemes [6]–[8] have been presented toward several directions and their experimental results have shown better coding efficiency than those of the conventional SBC scheme. Similarities among multiresolution subimages have been fully utilized in the fractal-coder and efficiently exploited for coding of images. In low bit-rate applications, however, the fractal-based SBC scheme causes block discontinuities due to its block-based processing, which makes it difficult to obtain subjectively good image quality.

This paper proposes a novel subband image coding scheme which exploits similarities among the multiresolution subimages while coding the edge information separately. In our scheme, an input image is separately coded in two sources; a region source such as a region boundary and a residue source such as texture. The region boundary is coded by contour coder while the residue source is coded by fractal-based subband coder. The residue source is independent of the region source and has a high-energy compaction, making it well suited for low bit-rate image coding. The proposed coding scheme is close to the human visual system because it is similar to human perception of the image that separates the region and the residue sources [5].

This paper is organized as follows. Section 2 summarizes the principle of the SBC scheme and introduces the energy compaction properties for the multiresolution subimages. Section 3 presents the proposed coding scheme including segmentation method for separating the region source from the input image. Section 4 includes the experimental results verifying the efficiency of the proposed coding scheme. Finally, Sect. 5 draws some conclusions.

Manuscript received October 31, 2000.
Manuscript revised April 9, 2001.
†The authors are with the Graduate School of Engineering, Tohoku University, Sendai-shi, 980–8579 Japan.
a) E-mail: wey@mk.ecei.tohoku.ac.jp
2. Subband Image Coding Theory

A subband decomposition for the still image has first been introduced by Woods and O’Neil [1], and several methods [2]–[4] have been developed using various choices of analysis/synthesis filters and encoding/decoding schemes. In this section, we will briefly introduce the principle of SBC and previous works. Also, we introduce the energy compaction property among the decomposition subimages.

2.1 Principle of SBC Scheme

Two-channel filterbanks for the SBC consist of two parts: an analysis part and a synthesis part. A general scheme for a two-dimensional subband decomposition is based on a row-column approach as shown in Fig.1, where \( H_0(z) \) and \( H_1(z) \) denote a low-pass filter and a high-pass filter, respectively. This is regarded as the first two levels in a filter tree structure. The first stage represents a row filtering and the second stage represents a column filtering. Each subimage \( F_{j}^{l} \) \( (j = \{ LL, HL, LH, HH \}) \) is obtained first by filtering the rows of the input image with \( H_l(z) \) \( (l = 0, 1) \) and down-sampling by a factor of two. Then, the rows of each down-sampled image are filtered with \( H_m(z) \) \( (m = 0, 1) \) and down-sampled again. For multi-level subband decomposition, this decomposition process is applied to \( F_{LL}^{1} \) to produce four more subimages. Figure 2 shows an example of three-level subband decomposition for an image. In the synthesis part, the reconstructed image \( \tilde{F} \) is obtained by up-sampling the columns of \( \tilde{F}_{k}^{l} \) and filtering with \( G_l(z) \). Then, the rows of the filtered image are up-sampled and filtered with \( G_m(z) \).

A perfect reconstruction (PR) of the input image \( F \) can be obtained by an appropriate design of the analysis filters \( H_l(z) \) and synthesis filters \( G_l(z) \). Although the PR condition is not necessary for lossy image coding, it is desirable because reconstruction errors can be attributed to controllable lossy operations such as quantization. A typical method for achieving the PR is to choose the analysis/synthesis filters’ coefficients as follows:

\[
\begin{align*}
g_0(n) &= h_0(-n) \\
g_1(n) &= h_1(-n) \\
h_1(n) &= (-1)^{1-n}h_0(1-n),
\end{align*}
\]

where \( h_0(n) \) is a filter of which \( z \)-transform \( H_0(z) \) satisfies the following equation

\[
H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2.
\]

Note that Eq. (2) implies that the filter \( h_0(n) \) and its even translation form an orthogonality, and that the two filters \( h_0(z) \) and \( h_1(z) \) are orthogonal [9]. For the analysis/synthesis filters, quadrature mirror filters (QMF), proposed by Croisier et al. [10], are widely used because of their simple filter design and good frequency characteristics as filters.

Subimages produced by the subband decomposition are separately coded using their statistical properties. For the lowest frequency subimage \( (F_{LL}^{k}) \), differential pulse code modulation (DPCM) coding and transform coding technique are generally used since the correlations between pixels in the lowest frequency subimage are similar to those of the input image. For the higher frequency subimages \( (F_{LH}^{k}, F_{HL}^{k}, F_{HH}^{k}) \), pulse code modulation (PCM) coding is used because of the small correlations between pixels in the subimage.

2.2 Energy Compaction Property for Edges

In this subsection, we first investigate an energy com-
paction property for edges which is included in the image. Next, we will consider how to design the encoder for obtaining subjectively good quality of the reconstructed image. To observe the energy compaction rate for edges, we have chosen the three-level subband decomposition, which is partitioned in the frequency domain. The corresponding subimages of a test image are shown in Fig. 3(a). We extract the edges of the \((i, j)\)-th subimage \(f_i^j\) \((i = 1, 2, 3\) and \(j = \{LH, HL, HH\}\)) to be an element of \(f_i^j\), where \((x)\) means the nearest integer value of \(x\) and \(2^i\) is the downsampling rates for the \((i, j)\)-th subimage, respectively. This step is performed recursively until all edges of \(F(m, n)\) are processed.

3. If \((m, n)\) is an edge point of \(f_i^j\), declared in step 2, and \(\max\{|m' - m|, |n' - n|\} = 0\) or 1, declare \((m', n')\) to be an element of \(f_i^j\). Figure 3(b) shows the resulting edge-region of the test image Lena. The energy compaction of the \(i\)-th subimages can be derived from [5]

\[
\xi_i = \sum_j \sum_{(m,n) \in f_i^j} |F_i^j(m,n)|^2.
\]  

(3)

The energy compaction ratio of the \(i\)-th subimages is defined by

\[
\gamma_i = \frac{\xi_i}{\sum_{i=1}^3 E_i^i},
\]

(4)

\[
E_i = \sum_j \sum_{(m,n)} |F_i^j(m,n)|^2,
\]

where \(\gamma_i\) denotes the ratio of energy compaction of the \(i\)-th subimages to the total high-frequency subimages. From Eq. (4), the Lena image contained 63.8\% of the total high-frequency energy in their edge-regions when the edge pixels detected in the original image were taken as the 100\%.

This experimental result leads to the following facts:

- Many edges of the input image spread over the higher frequency subimages through the subband decomposition.
- In the SBC scheme, the quantization errors occurring in encoding spread over the reconstructed-image, entirely.

The high bit-rate reduction in the conventional SBC scheme [1] inevitably leads to severe boundary artifacts which result in a rapidly deteriorating image quality.

In order to preserve the edge information at low bit-rate compression, we divide an input image into a region source and a residue source. Our coding scheme utilizes the region source as a priori information for the subsequent reconstruction. Comparing with the input image, the residue source, difference between the input image and the region source, achieves relatively good energy compaction into the lower frequency subimages. Subband decomposition of the residue source reveals strong similarity between subimages at different scales. Each block in the subimages can be predicted by an affine transformation of blocks in lower resolution subimages. The separate coding for the region and the residue sources is very much agreed with the human visual system.
3. Proposed Image Coding Scheme

The block diagram of the proposed coding scheme is shown in Fig. 4. The proposed image coder consists of three stages: image segmentation for region extraction of the input image, region source coding, and residue source coding by prediction of multiresolution subimages using the fractal-based coder. The detail of each stage is described in the following subsections.

3.1 Image Segmentation

The goal of this stage is to extract the region source from the input image. The region source is a set of homogeneous regions, which is described by the region boundary and the mean of pixel values in each region. In this stage, we employ a hierarchical segmentation algorithm based on the mathematical morphology [12], [13]. The segmentation is processed the following three steps:

Step 1. Image simplification: This step removes a smooth edge from the input image in order to make segmentation easier. In our scheme, morphological open-close filter is iteratively applied until a certain number of contours are left. The simplification process controls the number of contours.

Step 2. Feature extraction: This step detects the presence of homogeneous regions. The output of feature extraction can be a set of markers which identify the boundaries of the homogeneous region to be segmented. A marker is a set of connected pixels which have the same pixel value. The pixel value is determined as the mean of pixel values in the region.

Step 3. Contour decision: This step locates the boundaries among the regions extracted in Step 2. After Step 2, there may be pixels that are not assigned to any regions. Therefore, the regions are expanded or contracted so that each pixel is assigned to the only one region. The watershed algorithm [13] is used in this process.

Figure 5(b) shows the segmentation example for test image Lena (Fig. 5(a)).

3.2 Region Source Coding

The contours and the gray level values of each region are coded in this stage. For encoding the contours of each region, the JBIG standard method is used [14] since it shows better performance than the well-known chain-code-based methods [15]. The JBIG coder is an adaptive binary arithmetic coder that dynamically adapts to statistics for each template, which is composed of the immediate neighborhood pixels already coded. On the other hand, the gray-level values of each region are quantized and coded by the entropy coder.

3.3 Residue Source Coding

The residue source is the difference between the input image and the region source. Using SBC scheme, the residue source is split into subimages which are associated with different regions in the frequency domain. We consider the three-level subband decomposition of the residue source.
In this stage, blocks in \( F_i^{LH}, F_i^{HL}, \) and \( F_i^{HH} \) (\( i = 1, 2 \)) are predicted from those in \( F_{i+1}^{LH}, F_{i+1}^{HL}, \) and \( F_{i+1}^{HH} \), respectively. The lowest resolution subimages \( (F_3^{LL}, F_3^{LH}, F_3^{HL}, F_3^{HH}) \) are coded separately from the remaining subimages. This is because the visual quality of the reconstructed image strongly depends on the fidelity of the reconstructed lowest subimages. After all the subimages are predicted from the blocks of lower resolution subimages, an approximation of the residue source is reconstructed from coefficients in the subimage.

The relationship between the residue source coding and conventional fractal coding [16] is as follows. As shown in Fig. 6, each subimage \( F_i^{LH} \) (or \( F_i^{HL}, F_i^{HH} \)) is divided into a set of non-overlapping range blocks \( R = \{ r_i; \; i = 0, \ldots, m \} \) of which size varies according to the resolution of subimages. The pool of domain blocks \( D = \{ d_j; \; j = 0, \ldots, n \} \) for the image \( F_i^{LH} \) (or \( F_i^{HL}, F_i^{HH} \)) consists of the blocks of the same size extracted from the subimage \( F_i^{LH} \) (or \( F_i^{HL}, F_i^{HH} \)). For each range block \( r_i \) in the subimage \( F_i^{LH} \), we will find a best matched domain block \( d_j \) in the subimage \( F_i^{LH} \) such that the transformation \( T_i \) is defined as

\[
T_i(d_j) = s_i \mathcal{J}_i(d_j),
\]

where \( s_i \) is an amplitude scaling factor and \( \mathcal{J}_i \) is an isometry operator including rotations of the domain block. The transformations \( T_i \) are similar to those of fractal-based coder [16]. However, unlike the conventional fractal coders which use the mean value as the offset parameter, there is no offset parameter in our scheme. This is because the mean value is subtracted from residue source.

The transformation \( T_i \) is chosen to minimize the distance between the range block \( r_i \) and its approximation \( T_i(d_j) \). The distance is defined as the following equation,

\[
D(r_i, T_i(d_j)) = \frac{1}{r_i^2} \left[ \sum_{l=1}^{B_i} \sum_{m=1}^{B_i} (r_i(l, m) - T_i(d_j(l, m))) \right],
\]

where \( B_i \) denotes the pixel size of the range block \( r_i \) and \( (l, m) \) is the coordinate of the pixel in the block.

Once the domain block \( d_j \) and the optimal transformation are found, the distance \( D(r_i, T_i(d_j)) \) is compared with the pre-defined threshold value \( \theta \). If \( D(r_i, T_i(d_j)) \) exceeds the threshold \( \theta \), the range block is divided into four sub-size blocks like in the quad-tree scheme. When \( D(r_i, T_i(d_j)) \) is less than the threshold \( \theta \), the range block \( r_i \) is coded with the transformation \( T_i \) parameters. The code for each range block is given by the location of the best matched domain block and the parameters defining the transformation \( T_i \), i.e. the value of the scaling factor and the index of the rotation. This procedure is performed from lower resolution subimages to higher resolution subimages, recursively.

4. Experimental Results

We programmed our coding scheme in C language, and tested it on several standard images. The size of each image was 512 \( \times \) 512 with 8-bits per pixel (bpp). For region source, we applied the morphological open-close filter with the window size of 5 \( \times \) 5 pixels to original images. We took as markers all regions of size larger than 90 pixels. The threshold for merging a region was set to 50.0. The adaptive binary arithmetic coding [14] was used to encode the contour of each region. About 15\% of the total bit-rate was spent in coding the region source. In all our experiments, we used the the QMF of length 16 [10] as analysis and synthesis filters. For residue source, we used three-level subband decomposition with the lowest subimages having resolution 64 \( \times \) 64. We considered two possible range-block sizes, 8 \( \times \) 8 and 4 \( \times \) 4, at the subimages \( F_2^{LH}, F_2^{HL}, \) and \( F_2^{HH} \). For the subimages \( F_1^{LH}, F_1^{HL}, \) and \( F_1^{HH} \), two possible range-block sizes, 32 \( \times \) 32 and 16 \( \times \) 16, were used. In those cases, the threshold parameter \( \theta \) for splitting a block was set to 10.0. The subimages at the lowest resolution were coded independently using DPCM and PCM coder followed by normalized Laplacian quantizer with numbers of levels 2, 3, 4, 5, 8, and 32. For the coding of the other subimages, transformation parameters were coded using the scalar quantizer. The encoder output bit rate was measured in b/pixel (bpp), and, as a measure of the reconstruction image quality, we used the peak signal-to-noise ratio (PSNR), which is defined as

\[
\text{PSNR} = 10 \log_{10} \left( \frac{255 \times 255}{(1/MN) \sum_{i=1}^{M} \sum_{j=1}^{N} (F(i,j) - \hat{F}(i,j))^2} \right),
\]

where \( F \) and \( \hat{F} \) denote the input and the reconstructed images, respectively, and \( M \) and \( N \) are the vertical and
the horizontal dimensions of the image, respectively.

We compared the performance of the proposed coding scheme with those of the high-performance im-

Fig. 7  PSNR versus bit-rate comparison for the 512 × 512 image Lena.

age coding schemes by Rinaldo and Galvagno [6], by Davis [7], and by DCT-based JPEG. For the image Lena, various bit rates versus PSNR for different coding schemes are plotted in Fig. 7. The bit rate includes all the information necessary to decode the image. Experimental results show that the proposed scheme gave a better performance than other coding schemes, such as Rinaldo and Galvagno [6], and JPEG, over the entire range of bit rate. For example, our coding scheme outperforms JPEG by 1.27 dB to 1.75 dB. Although the PSNR of the proposed scheme is similar to or slightly lower than those of Davis’ scheme, the proposed coded-image appears to be subjectively good image quality at low bit-rate.

An example of the perceptual differences of the proposed, Davis [7], and JPEG images coded at 0.25 bpp is shown in Figs. 8(a)–(d). Compared with other two coding schemes, the reconstructed image by the proposed scheme has less blocking and boundary artifacts. Even though some “muddy” degradation can be detected in the proposed coded-image, global visual quality is much better than the other two images. In
this experiment, about 0.18 bpp was required for the residue source, while the region source was coded at 0.07 bpp.

As another example, we applied our coding scheme to the image Cameraman which is coded at 0.2 bpp. Figures 9(a)–(d) shows reconstructed images for several schemes, respectively. The similar observation can be made for this image. The proposed coding scheme is subjectively more convincing, since the objects are delimited by sharp boundaries and exhibit smooth surfaces not degraded by blocking artifacts. In this case, 0.15 bpp was required for the residue source, while the region source was coded at 0.05 bpp. The average encoding time for test images was about 60 seconds on a SUN Workstation with 143 MHz Ultra1 processor.

5. Conclusions

In this paper, we have proposed a novel image coding scheme based on separate coding of region and residue sources. In order to preserve the edges of the reconstructed image, the input image has been separated into two sources which were coded separately. The region source containing boundaries and the mean value of each region have been coded using the binary adaptive arithmetic coder as contour coder and uniform quantizer. The residue source has been first decomposed into multiresolution subimages and coded using the fractal-based coder. This can be achieved by using the similarities among the subimages. The coefficients of each higher subimage have been predicted from those of the lower resolution subimages. The predicted blocks have been simply coded by specifying the location of the domain block and the parameters of the affine transform. Experimental results have shown that, for low bit-rate compression, the reconstructed images using the proposed coding scheme have better image quality than those reconstructed images using coding schemes such as JPEG and another SBC scheme [6]. Although the PSNR of the proposed scheme is similar to or slightly lower than that of Davis’ scheme, the reconstructed image of our scheme is considered not to have the blurring artifact which is shown in the image reconstructed by Davis’ scheme. Especially, the advantage that the proposed coding scheme has over the
other coding schemes is the lack of both any blocking and boundary artifacts in the reconstructed image.

References


Ho-Cheon Wey received M.E. degree of electrical communication engineering from Tohoku University, Sendai, Japan, in 1998. He is currently working towards the D.E. degree at Tohoku University. His main interests and activities are in image processing, particularly image and video codings. He is a student member of the IEEE.

Masayuki Kawamata received B.E., M.E., and D.E. degrees in electronic engineering from Tohoku University, Sendai, Japan, in 1977, 1979, and 1982, respectively. He was an Associate Professor in the Graduate School of Information Sciences at Tohoku University and is currently a Professor in the Graduate School of Engineering at Tohoku University. His research interests include 1-D and multi-D digital signal processing, intelligent signal processing, fractals and chaos in signal processing, and linear system theory. Dr. Kawamata received the Outstanding Transactions Award from the Society of Instrument and Control Engineers of Japan in 1984 (with T. Higuchi), the Outstanding Literary Work Award from the Society of Instrument and Control Engineering of Japan in 1996 (with T. Higuchi) and the 11th IBM-Japan Scientific Award in Electronics in 1997. He is a member of the IEEE, the Society of Instrument and Control Engineers of Japan, the Information Processing Society of Japan, Robotics Society of Japan, and the Institute of Image Information and Television Engineers. He is an IEEE Senior Member.