An Embedded Image Coder Using A Pruned Tree Structure of Wavelet Coefficients

Joon Young Choi[1] and Yo-Sung Ho[2]

Electronics and Telecommunications Research Institute[1]
161 Kajong-dong Yusong-gu, Taejon, 305-350, Korea
Phone: +82-42-860-3886, Fax: +82-42-860-6403
E-mail: jyc@etri.re.kr

Kwangju Institute of Science and Technology (K-JIST)[2]
1 Oryong-dong Puk-gu, Kwangju, 500-712, Korea
Phone: +82-62-970-2211, Fax: +82-62-970-2204
E-mail: hoyo@kjist.ac.kr

Abstract: Wavelet transform for image processing is a form of subband coding, where the image frame is decomposed into frequency subbands hierarchically and recursively. A tree structure of wavelet coefficients allows us to predict important coefficients across the scales and therefore to compress their coordinates efficiently. In this paper, we combine the root node with its descendants that correspond to wavelet coefficients and search the tree in a depth first manner. With the depth first search, a tree can be treated independently and the tree pruning process can be performed to make an optimal tree shape in a rate-distortion sense without losing the embedded nature of the coder.

1. Introduction

With growing requirements for efficient representation of still pictures, Joint Photograph Experts Group (JPEG) has produced a coding standard, IS 109181, for still image compression. Although the current standard is successful and incorporated into a large number of software products, it fails to produce the best picture quality or performance in some application areas. Current JPEG activities are aiming at creating a new image coding system, JPEG 2000, to provide a enhanced image coding system including the improved performance of low bit rate compression.

For low bit rate image coding, the compression framework usually consists of three stages: (1) image transform, (2) quantization, and (3) entropy coding. Image coders using block-based transforms, such as DCT in the current JPEG standard, provide a fairly good image quality at medium and high bit rate [1]. However, when they are coupled with coarse quantization, the well known blocking artifacts occur because they do not consider the interaction between pixels in neighboring blocks. Wavelet transform techniques, on the other hand, do not suffer from blocking artifacts and typically produce better quality images at low bit rates.

Wavelet transforms for image processing is a form of subband coding where images are decomposed into frequency subbands hierarchically and recursively. Almost all the energy of wavelet transformed image is compacted into the large-scale subbands and very few coefficients of a high magnitude in the small-scale subbands. For efficient compression, the property of interband dependencies in the wavelet transformed image subbands is exploited to specify the locations of a small number of coefficients with large energy. The tree structure allows us to predict important coefficients across the scales and to compress their coordinates efficiently. The principal strategy of adopting the tree structure is full exploitation of the interband dependencies. Trees are constructed with the nodes that correspond to the wavelet coefficients, and partitioned into small trees in the process of comparing the nodes with a given threshold. This is a kind of quantization process with an embedded nature [3]. In the embedded coder, a sequence of decreasing threshold values is applied to determine the significance, and the compression processes are executed iteratively.

In this paper, we proposed a new algorithm in which a hierarchical tree is composed of coefficients with interband dependency and then encoded in an embedded way. The significance of a subtree as well as the significance of wavelet coefficients is considered in the proposed method.

2. Tree Structure of Wavelet Coefficients

Wavelet image decomposition can be thought of as a tree-structured set of wavelet coefficients, providing a hierarchical data structure for representing the image, with each coefficient corresponding to a spatial location in the image.

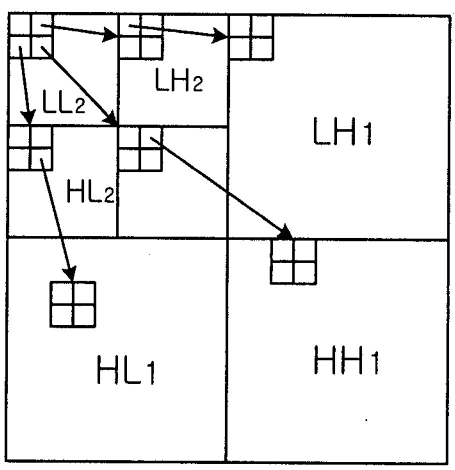


Figure 1. The parent children dependencies of subbands.

Figure 1 shows a two-level wavelet decomposition of an image. A spatial wavelet coefficient tree is defined as the set of coefficients from different frequency bands that represent the same spatial region in the image. Arrows in Figure 1 identify the parent-children dependencies in the tree. The coefficient at the higher level in the tree is called the parent, and all coefficients at the lower level are called children. For a given parent, the set of coefficients at all lower levels of the tree is called descendants. Similarly, for a given child, the coefficients at all higher levels is called ancestors. With an exception of the lowest frequency subband, all parents have four children.

3. The EZW and SPIHT Algorithms

The embedded zerotree wavelet (EZW) coder, proposed by Shapiro, exploits the correlation that still exists among the wavelet coefficients to generate zerotrees and encode them using a significance map and residue data [3]. The significance map is a binary bit map that represents the significance of wavelet coefficients. The significance of a wavelet coefficient is decided by comparing the coefficient with an exponentially decreasing threshold value. A block of coefficients that is proven to be insignificant is coded with the zerotree.

The EZW algorithm improves coding performance over conventional wavelet coders, and produces embedded bit streams that are appropriate for progressive image transmission. An embedded code represents a sequence of binary decisions that distinguish an image from the null, or all gray, image. Since the embedded code contains all lower rate codes embedded at the beginning of the bit stream effectively, bits are ordered in their importance. Using the embedded code, the encoder can terminate the encoding operation at any point; thereby, allowing the target rate or the distortion metric to be met exactly.

The set partition in a hierarchical tree (SPIHT) algorithm, proposed by Said and Pearlman, provides even better performance than the EZW algorithm by encoding the insignificant blocks of wavelet transform coefficients with smaller data bits [4]. With the same way as EZW, the SPIHT algorithm exploits the correlation between the subbands that are resulted from wavelet transform, and generates an embedded bit stream. The significance of each coefficient is also determined by the threshold operation. The SPIHT algorithm proceeds the encoding process, however, by partitioning coefficient sets and treating insignificant coefficient blocks in a different way. It puts coefficient sets in the predefined lists following the encoding order and defines a 'B' set to encode insignificant coefficient blocks efficiently.

4. The Pruned Tree Embedded Coder (PTEC)

4.1. Introduction

The embedded coders, such as EZW and SPIHT, utilize the zerotree to encode the position information of significant coefficients that are occurring rarely. If a coefficient in a higher level is coded with zero, the coefficients in the lower levels with the same orientation and location are combined with a tree and coded on the assumption that they are also insignificant. However, this assumption is not always true. A lower-level coefficient can have a larger magnitude than a higher-level coefficient. In such a case, there can be some significant coefficients that are not quantized to zero, whereas their ancestors, i.e., coefficients in the higher level, are insignificant. The insignificant coefficient with significant descendants is defined as an 'isolated zero' in EZW.

Consider a hierarchical tree where every node corresponds to the wavelet coefficient with same orientation and location (see Figure 2). When an isolated zero occurs, many positions of the insignificant nodes also should be specified individually to make the position of significant coefficient clear. This results in degradation of the coding efficiency. In the space-frequency quantizer (SFQ) [5], the hierarchical wavelet coefficient tree is pruned in the rate-distortion sense to overcome the inefficiency of the embedded coders. Although achieving marginally better coding efficiency, SFQ loses the embedding property that is very useful in many applications.

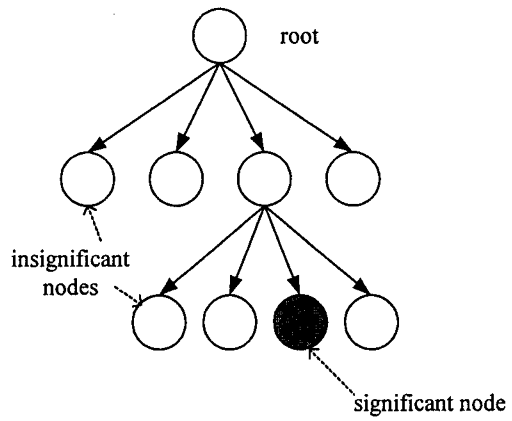


Figure 2. A hierarchical tree of wavelet coefficient

The pruned tree embedded coder (PTEC) improves coding efficiency by retaining a desirable tree shape based on the rate-distortion sense and produces an embedded bit stream. In order to maintain such characteristics, we arrange two kinds of coefficient significance test criteria:

(1) Significance of a coefficient by comparing it with a decreasing threshold T_k.
(2) Significance of a subtree where the coefficient resides.

The subtree is a branch of the hierarchical tree and consists of descendants with a root node. The root of the subtree can also be a descendant node of a tree or a subtree. The second criterion, i.e., significance of a subtree, is determined by means of the rate-distortion slope. The details are explained in the next section.

We also adopt a different way of tree searching in which a tree (or a subtree) is searched in a depth first manner. In the depth first searching, the significant test for a subtree is possible because a subtree can be treated independently.

In the decoding process, we also use threshold values to determine the significance of each coefficient and to reconstruct the coefficient. Threshold has a relationship of $T_k = T_{k-1}/2$ in each stage, and initial threshold T_0 is defined as follows [3]:

$$|x_i| < 2T_0 \quad (1)$$

where x_i is any coefficient in the frequency domain.

Although the initial thresold value can have any value, stisfying Equation (1), it is a power of two, in general [3][4].

4.2. Significance of the subtree

The following definitions of the tree and the subtree are used to decide significance of the subtree, i.e., the second criterion in Section 4.1.

(1) Z indicates a tree in the decoding process. Its root node P is in the 'tree list'.

(2) S^i is a subtree of Z with the root node, w^i, which is a member of the tree Z. The superposition i is an index of the root node and also applied to indicate a subtree.

(3) O^i is a children node set of S^i.

For the second criterion of the significant test, the slope between the expected values of distortion decrease D_k and rate increase R_k occurred by encoding each coefficient in the k-th stage are used. D_k and R_k be calculated as

$$D_k = p \cdot (1.5T_k)^k \tag{2}$$

$$\begin{aligned} R_k &= -(1-p)\cdot \log_2(1-p) - p \cdot \log_2 p + p \\ &= p + H(p) \end{aligned} \tag{3}$$

where p is the probability of occurrence of pixels that are greater than threshold. From Eq. (2) and Eq. (3), we can obtain a rate-distortion slope λ_k that is used as a significance test criterion of the subtree in the k-th coding stage.

$$\lambda_k = \frac{D_k}{R_k} = \frac{2.25 \cdot T_k^2}{H(p)/p+1} \tag{4}$$

The pruning parameter μ^i for each given subtree S^i is defined as the rate-distortion slope:

$$\mu^i \equiv \frac{d(S^i)}{d(S^i)} \tag{5}$$

where $d(S^i)$ indicates the cost of distortion by pruning S^i and $r(S^i)$ represents the gain of the bit rate. If the parameter value μ^i is smaller than λ_k, the subtree S^i is determined to be insignificant and coded as a zerotree. In other words, it is 'pruned'. $d(S^i)$ is defined as follows.

$$d(S^i) \equiv \sum_{j \in O^i} \left[\varepsilon^i + d(S^j)\right] \tag{6}$$

where

$$\begin{cases} (w^i)^2 & \text{if } w^i \text{ is significant} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Here, $d(S^i)$ is the sum of squared errors, resulting from elimination of coefficients which are larger than threshold.

$r(S^i)$ can be expressed as follows.

$$r(S^i) \equiv n(Z) - n(Z - S^i) \tag{8}$$

where $n(\bullet)$ indicates the number of bits generated in the process of encoding a given shape of the tree. It can be obtained by mimicking the encoding procedure. As explained in the previous section, depth-first searching for a tree make it possible to find out the number of bits generated to encode the subtree. $Z - S^i$ means the resulting tree after pruning the subtree S^i.

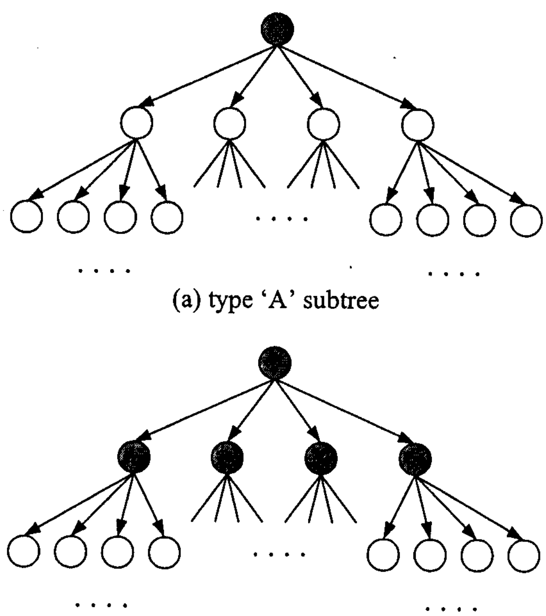
4.3. Encoding algorithm

In the coding process, a significance map is stored in the form of two linked lists. One of them is for the coefficients that are considered to be significant. It is called significant list (SL). Another list where the root node is combined with its descendant nodes is called the tree list (TL).

Entries of TL are compared with the current threshold. If an entry is found to be significant and it is a root node, the position of the root node is registered in SL. If the subtree sets are found to be significant, the tree is removed from TL and partitioned in a way that all children nodes become roots of the new subtrees. The partitioned subtrees are also tested for their significance and encoded in the same way as their parent tree. This procedure is repeated recursively until all the significance nodes in the tree are identified.

All the entries of SL are refined with respect to the current threshold except for those added during the current pass. The reason to skip these recently added coefficients during refinement is to equalize the maximum distortion for all coefficients, already encoded and still considered insignificant.

In order to make compression more efficient, we introduce two types of subtrees, as shown in Figure 3, in the same fashion as SPIHT. Type 'A' is a tree whose descendant nodes are identified as insignificant in the previous significance maps. Type 'B' is a tree that more than one of children nodes (direct descendants) are signified in the previous significance maps. With the two types of subtrees, we can prevent the subtree from being expanded into a few new subtrees for some time, and thus reduce the bit rate.



(a) type 'A' subtree

(b) type 'B' subtree

Figure 3. Two types of subtrees

The pruning and encoding operations of the PTEC algorithm are explained in Figure 4.

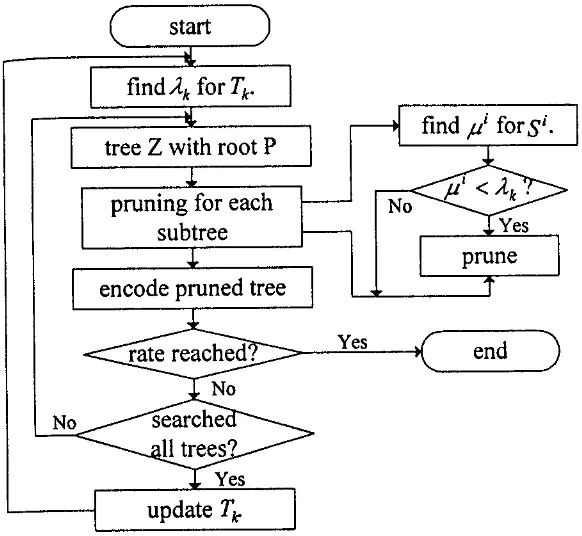


Figure 4. Block diagram of the PTEC algorithm

5. Entropy Coding

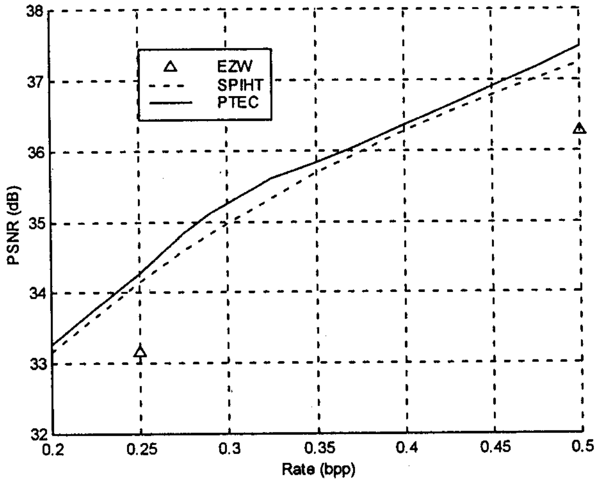
As with any other coding method, the efficiency of tree structured wavelet coder can be improved by entropy coding of the output. Although the refinement and sign symbols are approximately equilibrium between '0' and '1', the significance symbol is highly biased toward '0' and correlated with the significant status of its neighbors. This correlation can be exploited by entropy coding. We employ the QM coder, i.e., a binary arithmetic coder, as in JPEG standard [1]. In order to increase the coding efficiency, the arithmetic coder uses different probability tables for different contexts. Each adaptive model contains a probability distribution conditioned to the fact that a certain number of adjacent pixels are significant or insignificant. The dependency between magnitudes of adjacent pixels is fully exploited in this way. We used five contexts: four neighbor pixels (upper, down, left and right) and a pixel of the same location, but different orientation, called sibling.

6. Experimental Results and Conclusion

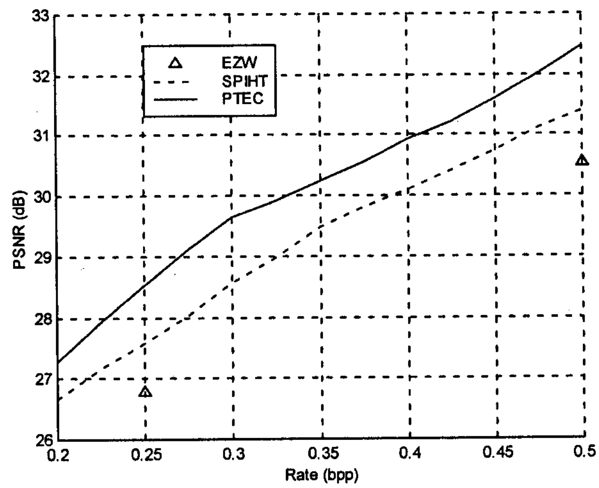
Performance comparison of PTEC with SPIHT and EZW is shown in Figure 6. The experiment is performed over monochrome images, Lenna and Barbara, of size 512×512 with 8 bpp. We adopt the 9/7 bi-orthogonal linear phase filter to perform the wavelet transform. As we can see in Figure 5(b), the proposed algorithm is more effective than EZW and SPIHT over Barbara. The improvement is, however, not so noticeable, compared to SPIHT, with Lenna that is simpler than Barbara. This is due to the fact that the isolated zeros occur frequently in complicated images and the pruning process plays an important role to obtain a good image compression.

We utilize PSNR as a measurement of distortion between the original and compressed images.

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right) \text{ dB} \tag{9}$$



(a) LENNA



(b) BARBARA

Figure 5. Performance comparison

References

[1] W. B. Pennebaker and J. L. Mitchell, *JEPG Still Image Data Compression Standard*, Van Nostrand Reinhold Publishers, New York 1993.

[2] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 674-693, 1989.

[3] J. M. Shapiro, "Embedded image coding using zerotree of wavelet coefficient," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.

[4] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 6, no. 3, pp. 243-250, June 1996.

[5] Z. Xiong, K. Ramchandran and M. T. Orchard, "Space-Frequency Quantization for Wavelet Image Coding," *IEEE Trans. Image processing*, vol. 6, no. 5, May 1997.