A Very Low Bit Rate Algorithm Based on Hierarchical Subband Scanning

[a]Humberto Ochoa, [a]Alejandra Mendoza, [a]Efrén Gutierrez, [a]Osslan Vergara, [b]Javier Vega, [a]Leticia Ortega, and [c]K.R. Rao

[a]Departamento de Ingeniería Eléctrica y Computación, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua, México.
[b]División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Chihuahua, Chihuahua, Chih., Mexico.
[c]Electrical Engineering Department, The University of Texas at Arlington, Arlington, Texas, USA.

Abstract: - A very low bit rate algorithm based on hierarchical partition of subbands in the wavelet domain is proposed. The algorithm uses the set partitioning technique to sort the transformed coefficients. The threshold of each subband is calculated and the subbands scanning sequence is determined by the magnitude of the thresholds which establish a hierarchical scanning not only for the set of coefficients with large magnitude, but also for the subbands. The algorithm is less complex computationally than SPIHT and JPEG2000. Results show that the present algorithm provides good image quality for very low bit rates.

Key-Words: - Embedded coding, wavelet image coding, image coding, progressive transmission.

1 Introduction

Many data compression techniques have been employed to decrease the amount of data required to represent information. For example, one data compression technique for image coding is the ISO/JPEG (International Standards Organization/Joint Photographic Experts Group) data compression standard which is based on the Discrete Cosine Transform (DCT). Although the ISO/JPEG technique has been widely adopted, its performance is not optimal [1].

Recently, new techniques based on wavelet transforms have emerged. These techniques achieve a hierarchical subband decomposition of a source image and the resulting transformed coefficients are transmitted using a progressive transmission scheme [2], [3], [4], [5]. Encoders like EZW [6], SPIHT [3], and SPECK [7] make use of this transmission scheme, which exploits the self-similarity of transformed coefficients across different scales.

SPIHT and SPECK maintain lists of significance so that the significant bits can be transmitted in decreasing bit plane order. This scheme makes the bit stream embedded. EZW does not perform the encoding of the coefficients; it only arranges the coefficients by its magnitude order. A lossless encoder, such as arithmetic or Huffman, is essential to compress the ordering information as conveyed by the results of the significance tests [3], [7].

2 Discussion

Suppose that the highest threshold of an image of size *MxN*, transformed into the wavelet domain using L levels of decomposition, is represented with k bits, and that the threshold is in the lowest resolution, lowest frequency subband (LL_L). Furthermore, suppose that the transformed image is going to be coded using the SPIHT coder and the threshold of each independent subband is calculated. Assuming the three subbands LH_L, HL_L and HH_L with same resolution as LL_L have a threshold represented by d_l bits, and the distance in bits from k to d_l is one, then, k is the first significant threshold and d_l is the next significant threshold. The total number of coefficients in LL_L is $\frac{1}{2^{2L}}(MxN)$. The number of sets containing 2x2 adjacent samples in each of the next LH_L, HL_L and HH_L subbands is $\frac{1}{2^{2(1+L)}}(MxN)$. Therefore, the sets in LH_L, HL_L and HH_L are insignificant to 2^k . To signal this condition a total of $\frac{3}{2^{2(1+L)}}(MxN)$ extra zeros must be sent to the decoder in the first sorting pass. If L is small, the subbands will have a higher resolution and the number of zeros sent will increase. Approximately $\frac{3(k-d_1)}{2^{2(1+L)}}(MxN)$ extra zeros must be sent to the decoder before a significant set can be found, while $k-d_1$ significant passes and $k-d_1-1$ refinement passes have been performed in LL_L subband; $k-d_1$ represents the distance in bits from the highest threshold to the next significant threshold.

Samples of the three subbands, at the next resolution level $L-1$, are grouped into sets of 4x4 adjacent samples. If the subbands have the next significant threshold and this threshold is represented by d_2 bits, the distance in bits from the previous significant threshold d_1 to d_2 is $(d_1 - d_2)$ bits. The number of zeros sent to the decoder to signal that the 4x4 sets are not significant is $\frac{3(d_1-d_2)}{2^{2(L-1)}}(MxN)$, and so on. If the distance in bits, between consecutives thresholds is large, the number of transmitted zeros increases. These zeros are part of a necessary header but do not convey any information related to a specific sample. For very low bit rates (rates less than 0.5 bpp) this condition may increase the rate without decreasing the MSE of the decoded image.

If the threshold of each independent subband is calculated and sent to the decoder, there is no need for sending extra zeros to signal the insignificant sets of consecutive subbands. Furthermore, the method of scanning each subband can be established by the magnitude of its associated threshold.

3 Coding methodology

The transformed image X has a hierarchical pyramidal structure. The highest frequency subband has the finest samples and lies at the bottom left of the pyramid. The lowest frequency subband has the coarsest samples and lies at the top right of the pyramid. We use the term coefficient for samples in the transformed domain represented with up to 16 bits. The transformed coefficient $cs_n(i,j)$ corresponds to a sample in subband n at position *(i,j)*.

Initially, the threshold of each subband is calculated. The subbands scanning sequence is established by the magnitude of the thresholds, subbands with large threshold are scanned first. The subbands are partitioned into sets of type S_q (where $q = 1....L+1$) of *mxn* adjacent coefficients, according to the subband they belong to. The resolution of a set in a finer subband is four times the resolution of a set in the next coarser subband. For example, the resolution of a type of set S_L is four times the resolution of a type of set S_{L+1}. In other words, the type of set S_L is made up four types of sets S_{L+1}. The type of set S_{L+1} consists of one coefficient only. The type of a set is the same if it belongs to a subband with the same resolution, and different if the set belongs to a subband with a different resolution, except for the coarsest subband whose set size is of one sample only. The type of sets S_3, S_2 and S_1 are illustrated in Fig. 1 for a two levels (L=2) of subband decomposition, using the same grouping methodology as SPIHT and SPECK [3], [7]. The difference is that we don't use spatial oriented trees but sets of type S_q.

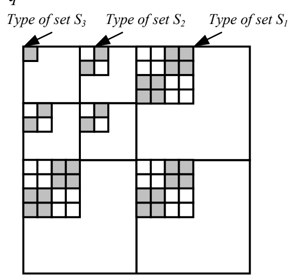

Fig. 1. Partitioning of a transformed image into sets of type S_3, S_2 and S_1.

Subbands have a maximum threshold of kn bits (where n = *1,2, ... (3L-1)*) represented by

$$kn_{(\max)} = \left\lfloor \log_2\left(\max|cs_n(i,j)|\right)_{(i,j)\in X|cs_n(i,j)|} \right\rfloor \quad (1)$$

A set is significant to a threshold kn if

$$2^{kn+1} > \max_{(i,j)\in S_q}\{|cs_n(i,j)|\} \geq 2^{kn} \quad (2)$$

After each sorting pass, the threshold is decremented by one and the subband scanning is repeated. The new threshold is compared against the remaining thresholds. If two or more thresholds are equal then, the subband or subbands with this threshold are included in the scanning process. When two or more thresholds are equal, the correspondent significant subbands are scanned in zigzag order from coarser to finer. If no threshold is equal, a new pass is performed over the previous significant subbands, but no new subband will be added to the scanning process. We say a subband is significant when the current threshold is equal to the threshold of the correspondent subband. Therefore, once the thresholds are known, there is no need for sending extra zeros to indicate that the sets are insignificant. Significant sets are split into four subsets and treated as one-fourth resolution sets in the next pass. The scanning process establishes a hierarchical order not only for sets of coefficients but also for subbands.

The significance of a set inside a subband can be written as [3].

$$\gamma(S_q) = \begin{cases} 1, & 2^{kn+1} > \max_{(i,j)\in Sq} \{|cs_n(i,j)|\} \geq 2^{kn} \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

Significant sets of type S_{L+1} are coded along with a sign bit. The sign bit can be expressed as:

$$\lambda c = \begin{cases} 1, & cs_n(i,j)|_{(i,j)\in S_{L+1}} < 0 \\ 0, & cs_n(i,j)|_{(i,j)\in S_{L+1}} \geq 0 \end{cases} \quad (4)$$

Refinement is performed on significant S_{L+1} sets from previous passes, according to the current kn^{th} threshold. The refinement of a previous significant sample is performed by extracting the kn^{th} bit of the coefficient.

Similar to SPECK, each set is partitioned into four subsets of one-fourth the resolution of the set [7]. However, the difference is that the size of the sets and the subband scanning sequence are defined at the beginning of the encoding process. The four subsets are compared against the threshold and if they are not significant, a zero is sent to the decoder to indicate the whole set is insignificant, if at least one subset is significant, a one is sent to the decoder. A four-bit map is sent to the decoder to indicate the position of the significant subset or subsets. The significant subsets are treated as individual sets and partitioned into four subsets of one-fourth the resolution of the set. Each subset is tested for significance and a four-bit map is generated to indicate the position of the significant subsets. Non significant subsets will be treated as sets of one fourth the resolution of the original set, in the next sorting pass. The process is repeated every time a new threshold is compared until sets of type S_{L+1} are reached.

4 Coding algorithm

The algorithm maintains two lists, one of significant sets (LSS) and one of insignificant sets (LIS). LSS has the upper left position (i,j) of sets that have been significant, while LIS contains the upper left position (i,j) of sets that have not been significant yet. The lists are initialized empty because the encoder knows the position of the sets tested based on the scanning sequence. The algorithm is applied to a rectangular image X which has been transformed into the wavelet domain with L levels of subband decomposition [8]. Image X contains the transformed coefficients $cs_n(i,j)$ at position (i, j) inside subband n. X exhibits a hierarchical pyramidal structure with the upper – left subband as the root of the pyramid. The algorithm starts by defining the sets in the coarsest subband as type of sets S_{L+1}. Each set contains only one coefficient. The sets in the finer subbands are defined as sets of type S_L ,S_{L-1}..... S_2, S_1. A set inside a subband is tested for significance against a threshold; if it is significant, its position is added to LSS and a one is sent to the decoder, otherwise the position of the set is added to LIS and a cero is sent to the decoder. If any of the sets of type S_L , S_{L-1}.... S_2, S_1 is significant, the set is quadrisected to one-quarter its resolution. Each subset is tested for significance using the current threshold (kn) and a four-bits map is constructed to indicate the position of the significant subsets. If a subset is not significant a cero is added to the map and its position added to the LIS. If the subset is significant a one is added to the map and its position added to LSS. The significant subset is treated as a set of one quarter the resolution of the original set. This new significant set is tested again and another four-bit map is obtained for each subset of one quarter of resolution. The process is repeated recursively until the set being tested is of type S_{L+1}.

5 Results

Coding method	**PSNR (dB)**			
	0.0017 bpp	0.06 bpp	0.26 bpp	0.48 bpp
SPIHT	7.33	19.86	22.61	24.60
JP2K	15.7	20.38	22.96	24.81
Proposed algorithm	9.71	20.39	22.75	24.39

Table 1. Comparison at very low bit rates.

Fig. 2. Coding results of Mandrill image at 0.06 bpp: original (top), SPIHT(middle), Proposed algorithm (bottom).

We presented the results of the standard monochrome, 8 bpp, 512 x 512 mandrill image. We used 5 levels of wavelet decomposition with 9/7 biorthogonal filters [8]. The different bit rates are calculated from a single encoded file. Table 1 shows the results obtained by this method, at very low bit rates, for the same image using SPIHT, JPEG 2000 VM 4.0 and the current algorithm. Fig. 2 shows the visual comparative results for Mandrill at 0.06 bpp. PSNRs can be seen in Table 1.

6 Conclusion

We introduce a new set partitioning algorithm, for very low bit rate, based on hierarchical scanning of frequency bands. The visual quality of the recovered images is better in the proposed algorithm than in SPIHT [3]. The key of the algorithm is to calculate the initial thresholds for each frequency band and use a progressive transmission scheme on the hierarchical pyramidal structure. Thresholds calculation avoids sending extra zeros at very low bit rates. The significance tests allow partitioning the set into subsets to group high energy areas. Different implementations can be performed; for example, the algorithm can be implemented to perform only the sorting pass and suppress the refinement pass section. The code is completely embedded and the computational complexity is low.

References:

[1] ISO/IEC IS 10918-1 | ITU-T Rec. T.81.

[2] Ochoa H. and Rao K.R., "A New Modified Hybrid DCT-SVD Coding System for Color Images," WSEAS Transaction on Circuits and Systems, vol. 4, pp. 1246-1253, Octubre 2005.

[3] A. Said and W.A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technology*, vol 6, pp. 243-250, June 1996.

[4] J. Andrew, "A simple and efficient hierarchical image coder," *Proc. IEEE Intl. Conf. Image Proc. (ICIP)*, vol. 3, Oct. 1997, pp. 658-661.

[5] ISO/IEC 15444-13|ITU-T Rec. T. 812.

[6] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3462, Dec. 1993.

[7] W. A. Pearlman, A. Islam, N. Nagaraj, and A. Said "Efficient, Low-Complexity Image Coding With a Set-Partitioning Embedded Block Coder," *IEEE Trans. Circuits Syst. Video Technology*, vol 14, pp. 1219-1235, Nov. 2004.

[8] M. Antonini, M. Barlaud, P. Mathieu, and I Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205-220, Apr. 1992.