REGION OF INTEREST CODING USING PARTIAL-SPIHT

Ahmed Abu-Hajar[a] *and Ravi Sankar*[b]

[a]Digital Video Research and Development, Pinellas Park, Florida
[b]Department of Electrical Engineering, University of South Florida, Tampa Florida
{abuhaja@eng.usf.edu, sankar@eng.usf.edu}

ABSTRACT

In this paper, a region of interest coding using partial-SPIHT (P-SPIHT) is proposed. P-SPIHT evaluates the probability of the significant coefficients *(P_1)* in each bit plane. Then it codes each bit plane independently and according to its P_1. The proposed algorithm codes the ROI region first and then it codes the background independently. The algorithm uses integer-to-integer shape adaptive discrete wavelet transform (ISA-DWT) which has the flexibility of transforming any number of isolated arbitrary shape ROI regions. The results show that compression of the proposed algorithm is superior to the ROI-based SPIHT coder.

1. INTRODUCTION

Set partition in hierarchal tree (SPIHT) has excellent rate-distortion performance as it was proposed in [1,2]. SPIHT is a wavelet-based image coder which is an enhanced version of the embedded zero-tree wavelet coder (EZW) proposed in [3]. Both algorithms, SPIHT and EZW, exploit the relationship of the coefficients among the different subbands, and then the coefficients are coded using the zero-tree. The zero-tree coder, as in SPIHT, checks for the significance of all the descendants at each node of the tree within a particular bit plane. If all the descendants of a particular node are insignificant a one bit that reads “0” is coded; otherwise, a one bit that reads “1” followed by the four bits of the node. The problem with SPIHT is that the coding efficiency degrades as the probability of the significant coefficients (P_1) increases. In [4], partial-SPIHT (P-SPIHT) was proposed to use one of three modes to code each bit plane. A specific mode is selected if P_1 falls within a predetermined range. In [5,6], P-SPIHT is enhanced further by classifying the data into several categories. Classifying the data increases the redundancy of the coded data which improves the compression of the arithmetic coder.

In this paper, we are extending P-SPIHT to support region-of-interest (ROI) coding. ROI coding is useful for images that contain critical information where the critical information is concentrated in small regions of the images. Since the ROI regions contain the critical information, they must be coded losslessly. On the other hand, the non-ROI regions or the background are coded lossy at lower bit rates. The lossless coding of the arbitrary shape regions requires the need of the integer-to-integer shape adaptive discrete wavelet transform (ISA-DWT) similar to the one proposes in [7]. These transforms are flexible because they use 5/3 filter to transform the ROI regions for lossless coding and they use the 9/7 filter to transform the background for lossy coding.

This paper is organized as follows. Section 2 presents a review of P-SPIHT. Section 3 presents the integer-to-integer shape adaptive discrete wavelet transform, and Section 4 presents the proposed ROI-based P-SPIHT coder. Section 5 presents the results and discussions, and Section 6 presents the conclusion.

2. REVIEW OF P-SPIHT

Since the probability of the significant coefficients, P_1, increases at lower bit plane, the compression efficiency of SPIHT decreased as it continues to code lower bit planes [4]. In order to get around this problem, P-SPIHT was proposed to select one of three different coding modes that code each bit plane independently and according to its P_1. A particular mode is selected if P_1 falls within a predetermined range where range of P_1 is optimum each coding mode [4]. The three coding modes are:

Mode_1: This mode codes the bit plane using SPIHT as long as P_1 is less than 0.2.

Mode_2: This mode is selected if P_1 is between 0.2-0.3. Simply, it checks for the significance of the 2x2 sets only and it ignores the spatial correlation among the descendant coefficients. If all the 2x2 sets are insignificant it transmits “0” else it transmits “1” followed by the four bits of the set.

Mode_3: This mode is selected if the P_1 is less than 0.3. It codes the four bits of the 2x2 set without checking for the significance condition.

Despite the fact that Mode_3 does not provide any compression, it is more efficient than SPIHT as the former would inflate the data by up to 20% instead of compressing it. P-SPIHT provides higher compression than SPIHT at higher bit rates, but the compression of both becomes identical at low bit rates. This is because at low bit rates only the higher bit planes are coded, and then P-SPIHT would select Mode_1 which is identical to SPIHT. In [5,6], the P-SPIHT was enhanced by sorting the embedded bits into different categories such that the data within each category is highly correlated. Then each category is entropy coded with its own frequency model using adaptive arithmetic coder similar to the one in [8]. The embedded data are sorted into three main categories and they are listed as:

1. *Magnitude Bits (MB)*: This category contains the magnitude of the coefficients only. It is further sorted to *Insignificant Magnitude Bits (IMB)* and *Significant Magnitude Bits (SMB).* Since the coder and the decoder knows which coefficients have been coded as significant prior to this bit plane, it is capable on sorting the coefficients that have not been coded as significant prior to the current pass (*IMB*) and the bits that are being coded as significant prior to the current pass (*SMB*).
2. *Tree Bits (TB)*: This category contains the bits responsible for checking the significance of each set. and similar to the *MB,* it sorts the *TB* into to *Insignificant Tree Bits (ITB)* and *Significant Tree Bits (SMB).*
3. *Sign Bits (SB)*: This category contains the sign of the coefficient. The sign bits are coded once after coding the magnitude bit as a significant bit for the first time.

The categorized data are coded in an embedded manner as shown in Fig.1. The whole image is coded bit plane by bit plane and each bit plane is coded slot by slot. The slots are coded in a sequential order of ITB, STB, IMB, SMB then SB. Each slot contains the data and End of Slot Symbol (ESS). The ESS synchronizes the information which indicate the end of coding a particular category within a bit plane.

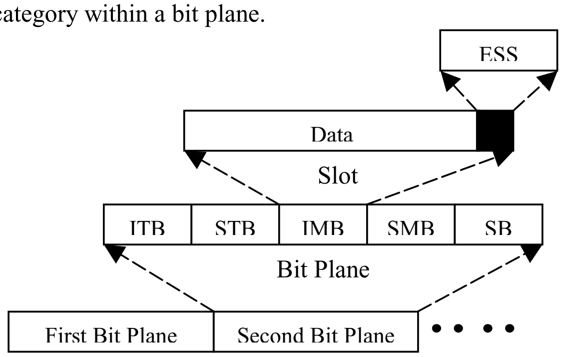

Fig. 1 Architecture of embedding the categorized data

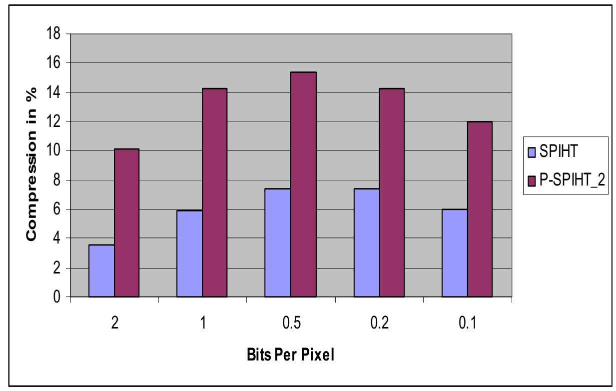

Fig. 2 The Average compression in % of the arithmetic coder for 8-bit images.

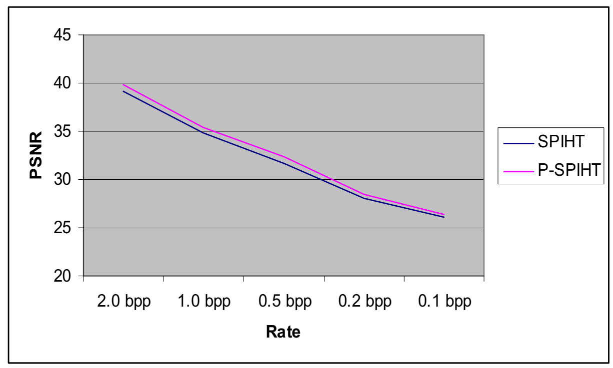

Fig. 3 PSNR at various rates for P-SPIHT and SPIHT

Sorting the data improves the performance of the arithmetic coder. Fig. 2 shows that sorting the data improves the compression of the arithmetic coder by a factor of two. Fig. 3 shows that P-SPIHT is superior to SPIHT for all bit rates.

3. INTEGER-TO-INTEGER SHAPE ADAPTIVE DISCERETE WAVELET TRANSFORM (ISA-DWT)

Integer-to-integer wavelet transforms, also called reversible transforms, have the advantages of providing lossless compression while using fixed-point arithmetic and entailing less complexity as they are implemented using the lifting scheme [9]. In [10], shape adaptive discrete wavelet transforms were proposed for conventional wavelet transforms, and then it was extended in [7] to support integer-to-integer shape adaptive discrete wavelet transforms (ISA-DWT). The transforms in [7,10] are capable on transforming several isolated ROI regions and the number of coefficients in the transform domain is equal to the number of pixels in the ROI regions. This allows the ROI region to be coded independently from the other background (non-ROI region), which adds the flexibility of using an optimum filter for each region and the bit rate. In this paper, ISA-DWT is used in order to support lossless compression and to reduce the arithmetic complexity.

ISA-DWT requires the prior knowledge of the arbitrary shape ROI region which is defined by a binary mask $M[n]$. It is possible to have several isolated regions that are defined by the binary mask where each isolated region is transformed separately from all others using ISA-DWT.

ISA-DWT requires the prior knowledge of the length (N) and the starting position (s) of the segment, and then the segment is symmetrically extended around its boundaries as shown in Fig. 4.

a_2	a_1	a_0	a_1	a_2	a_3	a_4	a_3	a_2

Fig.4 Symmetric extension of a signal

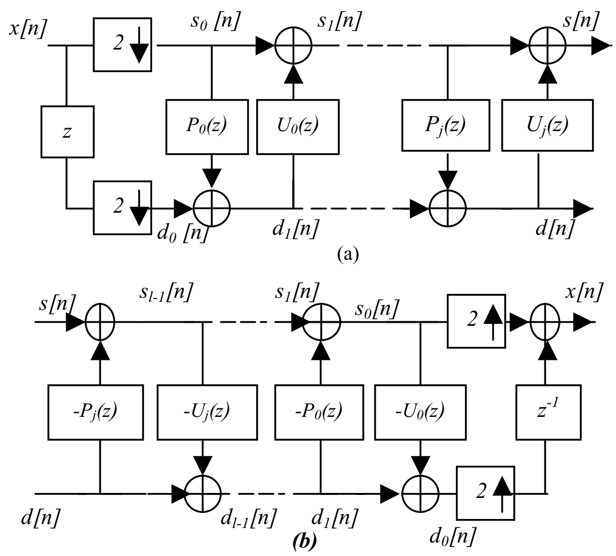

Fig. 5 (a) Lifting scheme of the forward transform (b) Lifting scheme of the inverse transform

Then, the signal is decomposed into the odd and the even indices of the signals as shown in Fig.5. The even and the odd signals start at $\left\lfloor \frac{s}{2} \right\rfloor$ where $\lfloor\ \rfloor$ is a truncation to the lowest nearest integer. ISA-DWT is evaluated according to the length of the segment as follows:

- If the length of the isolated segment is unity ($N = 1$), then the value of the pixel is appended to the low frequency coefficients, and the high frequency coefficient is marked as don't care.
- If N >1 and even then the transform is evaluated using lifting scheme as a regular signal.
- If N>1 and odd, then the first N-1 samples are evaluated using the lifting scheme and the last sample is append at the end of the low frequency coefficients.

In order to evaluate the transform for several levels, the original binary mask must be generated to new binary masks $M_s[n]$ and $M_d[n]$ for $s[n]$ and $d[n]$, respectively. The inverse of the transform is shown in Fig.5(b) and is evaluated similar to the forward transform.

The normalization of the coefficients which provides the same energy content of all the subbands within a particular bit plane is used for lossy coding and it is omitted for the lossless coding.

4. ROI-BASED P-SPIHT CODING

In this section, we are proposing to extend P-SPIHT to code any number of arbitrary regions. Despite the fact that the proposed ROI coder/decoder must know the location of ROI prior to coding, other techniques such as the maximum shift method may be used as well.

The architecture of the coder is shown in Fig. 6, which multiplies the image by its mask to generate the masked image. The ROI region is transformed using ISA-DWT. The mask must be generated in the transform domain to indicate the locations of the coefficients that belongs to ROI region. Then, the ROI-based P-SPIHT codes the decomposed ROI region on a bit plane basis.

The ROI-based P-SPIHT codes the coefficients that are part of the ROI region only. If the coefficient is not part of the ROI, the coefficient is discarded. Then, the coder evaluates the probability of the significant coefficients P_1 of each bit plane. Only, the coefficients that belong to ROI are used to evaluate the probability of each bit plane. Once P_1 is evaluated, the coder selects one of the coding modes presented in Section 2. Then the data are sorted into the five categories and each category is coded using adaptive arithmetic coder with its own frequency model.

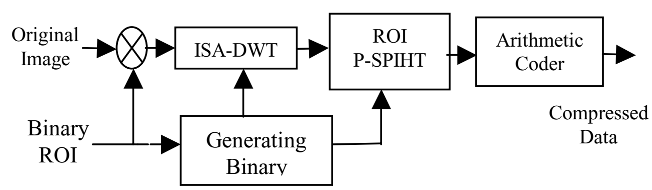

Fig. 6 ROI Coder Using P-SPIHT

5. EXPERIMENTAL RESULTS

Ten 8-bit 512x512 images were tested to evaluate the performance of the ROI-based P-SPIHT coder. In order to evaluate the coder effectively, the same ROI region is used for all the tested images. The ROI region has a circular shape and its center is located in the middle of the image, and it occupies about 12% of the total image area. The ROI region is coded using the 5/3 filter and the background is coded lossy using the 9/7 filter.

Table 1 shows that coding the ROI region losslessly using the proposed ROI-based P-SPIHT coder provides higher compression than ROI-based SPIHT by 10.7% on the average. The bit per pixel (bpp) for the arbitrary shape region is evaluated by the following:

$$bpp = \frac{total\ coded\ bits}{total\ bits\ within\ the\ region}$$

Table 2 shows that the lossy compression of the background using P-SPIHT is superior to SPIHT for all the bit rates. The bits rates of P-SPIHT are lower that SPIHT due to the efficiency of sorting the data, but they have the same image quality because P-SPIHT selects Mode_1 at low bit rate.

5. CONCLUSION

In this paper, ROI-based P-SPIHT was proposed which is capable of coding each arbitrary shape ROI regions independently. The compression of the proposed algorithm is superior to SPIHT for lossy as well as lossless coding.

References

[1] A. Said and W. Pearlman, "A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Trans. On Circuit and System for Video Technology,* vol. 6, no. 3, pp. 243-250, June 1996.

[2] A. Said and W. Pearlman, " An Image Multiresolution Representation for Lossless and Lossy Compression," *IEEE Trans. Image Processing,* vol. 5, no. 9, pp. 1303-1310, September 1996.

[3] M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelets Coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3462, December 1993.

[4] A. Abu-Hajar and R. Sankar, " Wavelet Based lossless Image Compression Using Partial SPIHT and bit plane Arithmetic coder," *ICASSP 2002* Orlando, Florida, May 2002.

[5] A. Abu-Hajar and R. Sankar, "Enhanced Partial SPIHT for Lossless and Lossy Image Compression," *Accepted ICASSP 2003* Hong Kong, April 2003.

[6] A. Abu-Hajar, "Wavelet-Based Region of Interest Coding for Medical Imaging," *Ph.D. Dissertation, Univ. of South Florida*, Tampa, November 2002.

[7] A. Abu-Hajar and R. Sankar, "Integer-to-Integer Shape Adaptive Discrete Wavelet Transform for Region of Interest Coding," *10th IEEE International DSP workshop*, Pine Mountain, GA, October 2002.

[8] I. Witten, R. Neal and J. Cleary, "Arithmetic Coding for Data Compression", *ACM Communication*, vol. 30, no. 6 pp. 520-540, June 1987.

[9] M. D. Adams and F. Kossentini, "Reversible Integer-to-Integer Wavelet Transforms for Image Compression: Performance Evaluation and Analysis," *IEEE Trans. Image Processing*, vol. 9, no. 6, pp. 1010 -1024, June 2000.

[10] S. Li and W. Li, "Shape-Adaptive Discrete Wavelet Transforms for Arbitrarily Shaped Visual Object Coding," *IEEE. Trans. Circuit and system for Video Technology*, vol. 10, no. 5, pp. 725-743, August, 2000.

Table 1. Evaluation of the ROI Lossless Compression

Image	ROI SPIHT (bpp)	ROI P-SPIHT (bpp)	Difference in (%)
Airplane	5.509	4.983	10.555
Lena	5.211	4.770	9.2452
Tiffany	4.759	4.273	11.373
Barbara	5.433	4.936	10.068
Baboon	5.554	4.981	11.503
Boat	5.717	5.179	10.388
Girls	4.898	4.419	10.839
Gold	6.059	5.436	11.460
Sky	7.452	6.672	11.690
Sailboat	5.437	4.950	9.8383
Average	5.603	5.059	10.731

Table 2. Evaluation of the Lossy (Background) Compression

Image	SPIHT		P-SPIHT	
	Bpp	PSNR	Bpp	PSNR
Lena	1.908	45.4887	1.752	45.6615
	0.930	40.0793	0.849	40.0793
	0.457	35.402	0.417	35.5402
	0.183	31.2773	0.168	31.2773
	0.091	28.6120	0.086	28.2773
Airplane	1.910	44.8139	1.757	44.9593
	0.930	39.9361	0.848	39.9361
	0.459	35.382	0.418	35.3820
	0.180	30.727	0.166	30.7270
	0.09	27.6516	0.085	27.6516
Baboon	1.938	31.6674	1.819	31.7622
	0.945	27.4502	0.872	27.4502
	0.463	24.5570	0.423	24.5570
	0.186	22.6971	0.172	22.6971
	0.092	21.6971	0.088	21.6971
Sailboat	1.916	37.9834	1.792	38.0738
	0.924	33.9613	0.841	33.9613
	0.442	30.7832	0.403	30.7832
	0.182	28.3521	0.170	28.3521
	0.087	25.1885	0.083	25.1855
Barbara	1.937	37.9989	1.787	38.0361
	0.946	31.4185	0.870	31.4185
	0.466	27.4012	0.434	27.4012
	0.184	24.4235	0.174	24.4235
	0.089	23.2932	0.085	2..2932
Average	1.9218	39.59046	1.7814	39.69858
	0.935	34.56908	0.856	34.56908
	0.4574	30.70508	0.419	30.73272
	0.183	27.4954	0.17	27.4954
	0.0898	25.28848	0.0854	21.02094