RATE DISTORTION OPTIMIZED VECTOR SPIHT FOR WAVELET IMAGE CODING

YongQing Liang and Scott E. Budge

Electrical & Computer Engineering Dept.
Utah State University, Logan, UT 84322
yongqing@cc.usu.edu, scott.budge@ece.usu.edu

ABSTRACT

In this paper, a novel image coding scheme using rate distortion optimized vector quantization of wavelet coefficients is presented. A SPIHT type algorithm is used to locate significant wavelet vectors, and they are classified into a number of classes based on their energies. The set partitioning bits are reused to indicate the vector classification indices, and a significant number of bits are saved. A set of codebooks with different sizes are designed for each class of vectors, and a Lagrangian optimization algorithm is employed to select an optimal one for each vector based on its operational rate distortion curve. The proposed coding scheme is capable of trading off between the number of bits used to code each vector and the corresponding distortion, and thus optimally allocating the bits among the vectors to minimize the total distortion. Experimental results show that our proposed method outperforms other zerotree-structured embedded wavelet coding schemes such as SPIHT and SFQ. A performance comparison with JPEG2000 is also presented.

1. EXTENDED SUMMARY

Wavelet based image coding has been shown to be an effective method. Space-frequency quantization (SFQ) [1], embedded zerotree wavelet (EZW) [2] and set partitioning in hierarchical trees (SPIHT) [3] are high-performance wavelet coders which benefit from using zerotree data types. A vector SPIHT method is proposed in [4]. The structure of the proposed rate-distortion optimized vector SPIHT (RDVSPIHT) and how it differs from the existing methods is presented in the followings.

There are four main characteristics of the proposed RDVSPIHT method. First, a vector set partitioning technique [5, 4] is used to locate the significant vectors. However, unlike SPIHT, refinement passes are not used since the wavelet coefficients can be recovered by mapping the indices to vectors through the corresponding energy codebooks. Second, the significant vectors are classified into a number of groups based on their energies to obtain relatively small vector spaces and thus the complexity of the vector quantization is reduced. Note that no extra bits are need to represent the classification indices because they can be indicated by the set partitioning information. Third, the codebooks are generated under an unconstrained condition given the sizes of the codebooks and their energy classes, and thus better performance is obtained with moderate computational complexity. Fourth, a rate distortion optimization algorithm is implemented to select a codebook with the minimum cost from a set of codebooks. In this way, the encoder can allocate bits to each coding unit based on its operational rate distortion curve and thus further improve the image quality. However, the codebook indices are required at the decoder side and this is the cost that the encoder needs to pay for rate distortion optimization. The codebook indices can be arithmetic coded to save some bits or sent directly to the decoder for less computation. If the gain of the rate distortion optimization is larger than the cost of coding the codebook indices, improvement is obtained.

Fig. 1 shows the block diagram of the proposed RDVSPIHT method. A discrete wavelet transform is employed on the original image to decorrelate the image pixels and compact the energy into low frequency and some spatially clustered high frequency coefficients. Then, blocks of the wavelet coefficients are grouped together to form vectors, and they are processed by a SPIHT type vector set partitioning coder. As in [3], there are a number of sorting passes to partition the vectors into significant and insignificant ones, and the significance information is stored in three ordered lists: list of insignificant sets (LIS), list of insignificant pixels (LIP), and list of significant pixels(LSP). In the i^{th} sorting pass, the energy of each vector in LIP and LIS is compared with a given threshold T_i, and its significance is described in an efficient way using zerotrees. After each pass, the threshold is reduced by half. If a vector becomes significant in the i^{th} sorting pass, its energy must be in the range $[T_i \; T_{i-1})$. Therefore, the bits used to denote the significance of the vectors can be reused to classify the vectors into a number of groups with different energy ranges. The vector space of each class is relatively small compared with original vector space and thus a moderate size codebook is possible for vector quantization to obtain desired performance. In addition, for each vector class, a

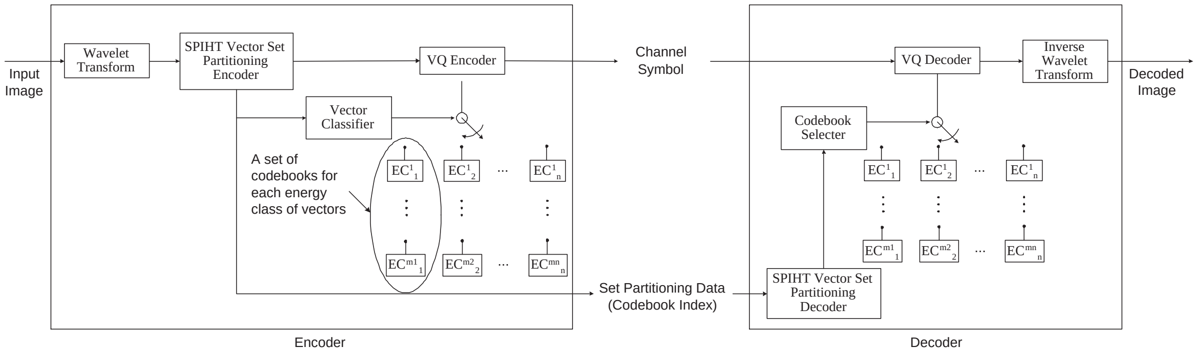

Fig. 1. Encoder and decoder structures of the proposed RDVSPIHT.

set of codebooks with different sizes were designed by a generalized Lloyd algorithm (GLA). A cost function is computed for each codebook, and the codebook with the minimum cost is selected as the optimal one through the formula given as,

$$n = \arg\min_{i \in S} \left(D(i) + \lambda R(i) \right) \tag{1}$$

where S is a set of codebooks for each energy vector class. In (1), n is the selected codebook index, which minimizes the total distortion given the Lagrangian multiplier λ.

Note that the solution to (1) is also a solution to the corresponding constraint problem, finding the optimal codebook resulting in the minimum distortion subject to the constraint of a target bit rate R_{target}. There exists a particular λ for each R_{target}, and the corresponding solution to (1) minimizes the distortion under the rate constraint. By varying λ from zero to infinity, the solution to (1) gives a convex operational R-D curve, and thus a bisection or a fast convex search can be used to find a particular λ to satisfy R_{target}.

The performance of the proposed RDVSPIHT method is demonstrated with natural grayscale images in this paper. For comparison, the results obtained by SPIHT and SFQ (implemented by QccPack [6]) are used as a reference. The results obtained by the CVSPIHT method [7] with fixed codebook size is also included to show the efficiency of the rate distortion optimization.

Table 1 shows the performance, in terms of PSNR (peak signal to noise ratio), of different image coding methods, including SFQ, SPIHT, CVSPIHT and the proposed RDVSPIHT, applied to the test images. Note that the images marked with "*" are not in the training set. In terms of PSNR, the proposed RDVSPIHT method performs consistently better than SPIHT and SFQ methods for the images both inside and outside of the training set. In addition, with rate distortion optimization, the proposed method is able to select the optimal codebook based on the cost function given by (1), and thus achieve higher PSNR compared with fixed codebook size CVSPIHT scheme. The final paper will also present a performance comparison with JPEG200.

Test images	Bit rate	SFQ	SPIHT	CVSPIHT	RDVSPIHT
Lena	0.228	30.24	30.77	31.62	32.31
Chair	0.370	29.58	30.16	30.91	31.55
Elaine	0.166	29.67	29.88	30.34	30.83
Goldhill	0.325	29.61	29.75	30.52	30.77
London *	0.379	30.24	30.34	31.91	32.03
Peppers	0.233	29.94	30.60	30.74	31.77
Boat	0.339	29.77	30.09	30.75	31.26
Brian *	0.207	30.68	30.95	31.43	31.88
Barbara *	0.463	28.52	28.95	29.60	29.90
anonymous	0.222	31.12	30.61	32.29	32.80

Table 1. PSNR values (in dB) and bit rates (in bits per pixel) of SFQ, SPIHT, CVSPIHT, and the proposed RDVSPIHT methods.

To examine the subjective image quality, a portion of the London image is shown in Fig. 2. The original image is shown in Fig. 2(a), while the encoded images with SPIHT, SFQ, and the proposed RDVSPIHT methods at bit rate of 0.379 bits per pixel are shown in Fig. 2(b)-(d), respectively. It can be seen that the proposed RDVSPIHT provides better visual quality than SPIHT, and SFQ such as better edges around the top of the towers and more details on the bridge.

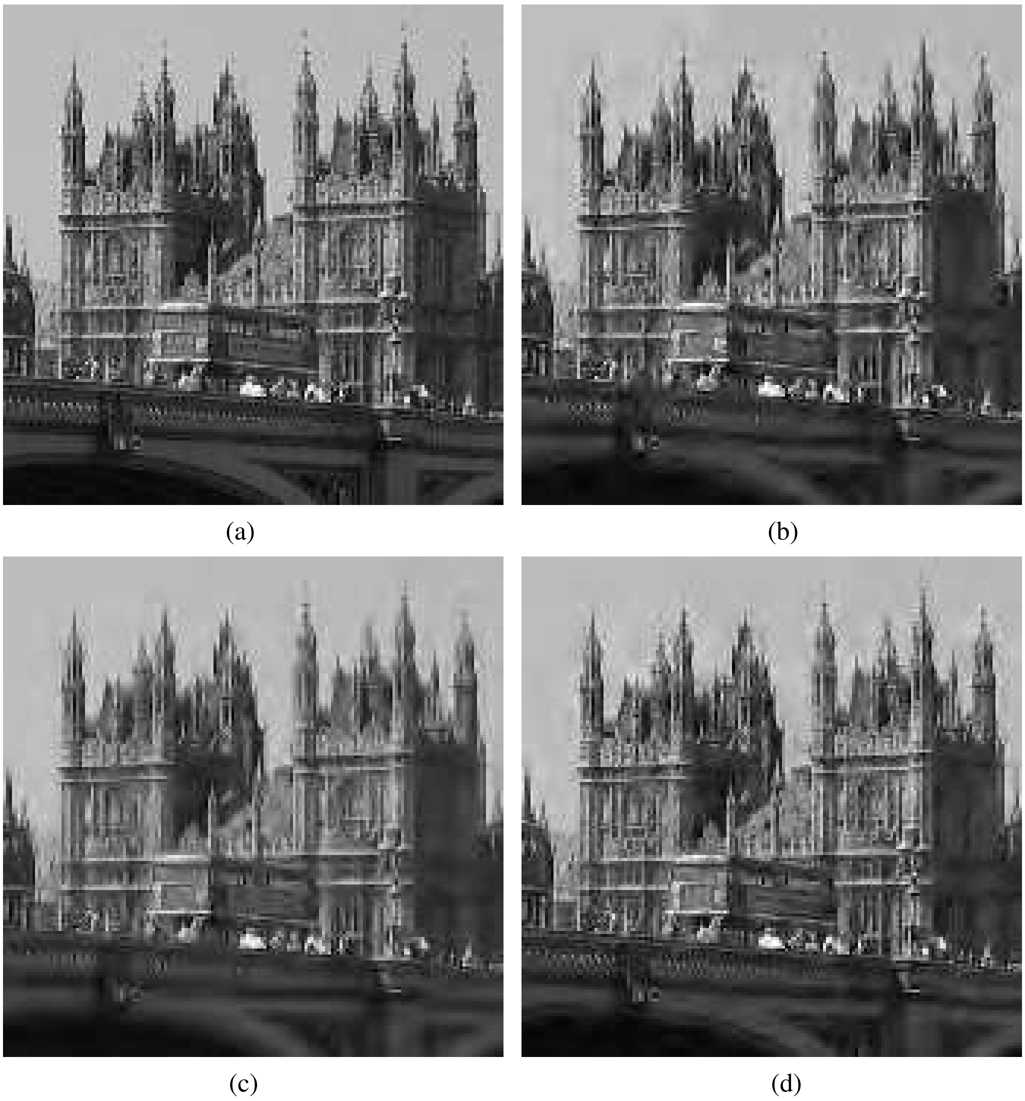

Fig. 2. Portion of the original and encoded London images at bit rate of 0.379 bpp. (a) Original, (b) SPIHT (PSNR = 30.34), (c) SFQ (PSNR = 30.24), and (d) RDVSPIHT (PSNR = 32.03).

2. REFERENCES

[1] Zixiang Xiong, K. Ramchandran, and M.T. Orchard, "Space frequency quantization for wavelet image coding," *IEEE Trans. Image Proc.*, vol. 6, no. 5, pp. 677–693, May 1997.

[2] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, pp. 3445–3462, December 1993.

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

[4] D. Mukherjee and S.K. Mitra, "Vector spiht for embedded wavelet video and image coding," *IEEE Trans. Cir. & Sys. for Video Tech.*, vol. 13, no. 3, pp. 231 – 246, March 2003.

[5] X. Li J. Knipe and B. Han, "An improved lattice vector quantization scheme for wavelet compression," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 239 – 243, January 1998.

[6] J.E. Fowler, "Qccpack: an open-source software library for quantization, compression, and coding," DCC, March 2000, p. 554.

[7] Yongqing Liang and Scott Budge, "Classified vector spiht for wavelet image coding," in *IEEE ICIP*, 2006.