CLASSIFIED 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

Wavelet based image coding has been shown to be an effective method, especially at low bit rate. Two successful approaches, space-frequency quantization (SFQ) and set partitioning in hierarchical trees (SPIHT) were proposed to predict insignificant wavelet coefficients. However, both of them employ a simple scalar quantization scheme instead of powerful vector quantization for coding of the significant wavelet coefficients. In this paper we propose a new image coding method that combines the SPIHT type set partitioning technique with classified vector quantization. That is, the set partitioning technique is used to locate significant wavelet coefficient vectors based on vectors' energies. In addition, the same information is used as classification indices to select the correct energy sub-codebook for vector quantization. The main idea of the proposed method is that the SPIHT type information is shared by set partitioning for locating significant vectors and classified vector quantization. In this way, many bits are saved for coding the side information required by classified vector quantization. Experimental results show that our proposed method achieves better performance, as compared to SPIHT and SFQ.

Index Terms— Vector quantization, Wavelet transforms

1. INTRODUCTION

Image coding has become an important element to reduce the storage and transmission time for digital images. Wavelet transform coding has been studied by many researchers. Space-frequency quantization (SFQ) [1] using wavelet zerotree [2] and set partitioning in hierarchical trees (SPIHT) [3] are two high-performance wavelet coders. The key feature of these algorithms is the efficient prediction of insignificant and significant wavelet coefficients. The significant coefficients are uniformly quantized followed by entropy coding in SFQ or bitplane coded in SPIHT, while the insignificant ones are discarded (or pruned). In the SPIHT coder, sorting bits (the bits used to locate significance coefficients) may be further compressed with arithmetic coding, while sign bits and refinement bits are not entropy coded since these are roughly uniformly distributed. However, both of the methods employ simple scalar quantization, and thus there are still some spatial redundancies among neighboring pixels. In order to exploit correlation between neighbor coefficients, vector quantization (VQ) can be considered to improve the coding performance.

Although VQ provides good performance, direct use of it suffers from a serious complexity barrier and hence it is limited to rather modest vector dimensions and codebook sizes for practical problems. Several techniques, including classified VQ (CVQ) and finite state VQ (FSVQ), have been developed which apply various constraints to the structure of the VQ codebook and yield a corresponding altered encoding algorithm and design technique. As a result, higher vector dimensions and larger codebook sizes become feasible. These methods generally compromise the performance achievable with unconstrained VQ, but often provide very useful and favorable trade-offs between performance and complexity [4].

In CVQ, a classifier is used to partition the codebook into subcodebooks. Please note that the similarity between the FSVQ and CVQ [4]. Both methods have multiple state codebooks, and the encoder switches among these codebooks based on the current state determined by previous coded vectors and the current input vector. The encoder sends the channel symbol of the minimum distortion codeword in the current state codebook and the decoded reproduction is determined by the combination of the channel symbol and the current state. Unlike CVQ, however, FSVQ is not permitted to send the additional bits or side information usually used to specify the state; the decoder must determine the state from the previous states and the channel symbols. The problem of FSVQ is that there may be a mismatch of state between the encoder and the decoder. In CVQ, side information is transmitted to ensure that both the encoder and decoder are in the same state. Of course, the total bits used to code the image is increased because of the coding of the side information. The benefit of our proposed method is that we use the set partitioning information to determine the state of the encoder and decoder, since both of them are based on the vectors' energies. Thus, no extra bits need to be sent. Furthermore, we can obtain a perfect match of states between the encoder and decoder. On the other hand, there are limitations in this scheme. The feature used to switch among the states and to partition the vectors must agree, that is, the vector energy used in this paper. If we want to add more features to determine the next state, such as side match FSVQ in [5], a next state function must be derived to implicitly represent the state used by the encoder and decoder.

In this paper, we propose a classified vector SPIHT (CVSPIHT) method for wavelet image coding. The same information is shared by set partitioning and vector classification. Compared with [6], we use SPIHT technique to efficiently locate the significant vectors. A vector SPIHT method is proposed in [7]. The main difference of our proposed method is that we use an unconstrained vector quantization method on each energy codebook instead of multistage vector quantization to achieve better quality. According to [4], multistage coders often have only two and occasionally three stages to maintain desirable signal to noise ratio. As we mentioned before, unconstrained vector quantization has good performance with high complexity. Since we use multiple sub-codebooks, moderate codebook size is possible, and complexity is not a large problem any more. Therefore, unconstrained vector quantization becomes a good choice of coding wavelet coefficients in a specified energy range.

The rest of the paper is organized as follows. In Section 2, some statistics of wavelet coefficients are shown to demonstrate the correlations of neighboring pixels. Section 3 provides the details of the proposed CVSPIHT method. Experimental results are presented in Section 4 and concluding remarks are given in Section 5.

2. STATISTICS OF WAVELET COEFFICIENTS

A good understanding of wavelet coefficient statistics is useful to increase the compression efficiency of images. In general, there are interband and intraband correlations in the wavelet coefficients. This is due to the fact that the wavelet transform compacts energy into low frequency coefficients, while also representing high-frequency energy in a few spatially clustered high-frequency coefficients. The former is the assumption used in SPIHT and wavelet zerotree, exploiting the self-similarity across scales with decaying energy. Efficient prediction of insignificant coefficients is the key of these successful approaches. On the other hand, some statistical models were proposed to model the intra band wavelet coefficients. In [8], Lam used a generalized Gaussian distribution to model the wavelet coefficients within a sub-band, and extracted the parameters of the distribution from the empirical data. An empirical description of vector classes was proposed in [7], and they suggested different codebooks could be used for each energy based vector class to further improve the coding quality. However, a complex vector quantization scheme was used for coding the significant vectors. In this paper, we group neighboring wavelet coefficients into vectors, employ SPIHT type set partitioning technique to predict insignificant vectors, and thus let the decoder know where the significant vectors are. In addition, we use a straightforward generalized lloyd algorithm (GLA) for vector quantizing the significant vectors to exploit correlations between neighbor coefficients in the same subband.

It is well known that vector quantization is superior to scalar quantization because vector quantization provides an increased freedom in choosing the partition geometry compared to the very restrictive geometry in the scalar quantization case [4]. Furthermore, vector quantization offers a great performance gain when the correlation between vector components is high. Since wavelet coefficients have spatial correlations, vector quantization is thus a good choice for coding of them. In addition, we observe that the higher the vectors' energies, the more correlations exist between the vector components.

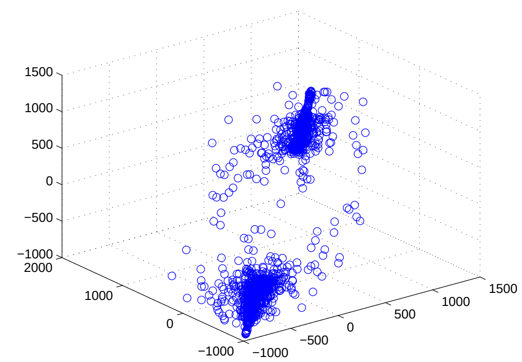

Fig. 1. The correlation of vector components with vector energy greater than 512.

Fig. 1 shows the correlations of the vector components when the vectors' energies are relative large, say greater than 512. The vectors are extracted from several test images which form the training set used in Section 4. The vector energy is just the L2 norm of the vectors. A 3-level wavelet decomposition with the 7/9 biorthogonal filters [9] is used, and the coefficients in each 2 × 2 block are grouped together to form a vector of dimension 4. Because it is impossible to plot points in a 4-dimensional space, to show the basic idea of the correlation of vector components, we just select the first 3 of them and give a 3-dimensional scatter plot. We can easily see that most of the points are clustered around the (1, 1, 1) line, and this shows the neighboring coefficients are heavily correlated.

3. CLASSIFIED VECTOR QUANTIZATION WITH SPIHT

Instead of coding scalar wavelet coefficients in SPIHT, groups of neighboring coefficients are formed into vectors, followed by classified vector quantization. To locate the significant vectors, a SPIHT type vector set partitioning coder [10] is used. However, unlike SPIHT, there is no refinement pass in our proposed method since the wavelet coefficients can be recovered by looking up the matched energy codebook. Because we use unconstrained vector quantization on each energy training set, better performance is obtained with moderate computational complexity. On the other hand, if a constrained vector quantization, such as TSVQ [4], is used on each energy set, a refinement pass can be added.

In general, the unconstrained method provides good quality, but it suffers from a serious complexity barrier, especially when the codebook is very large. Since we use multiple energy codebooks and moderate size codebooks are possible, unconstrained vector quantization scheme is adopted to achieve good performance. Note that, we still can obtain a embedded bit stream although unconstrained vector quantization is used. This is because we use SPIHT type information to transmit the vectors with relatively large energy first, and the decoder can recover the image based on the bits which it has received. The more bits it decodes, the better the quality that is obtained. In addition, the use of different size energy sub-codebooks is another way to adjust the bit rate of the coded image.

Fig. 2 shows the block diagram of the proposed CVSPIHT method. A discrete wavelet transform is employed on the original image to 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 combination of a SPIHT type vector set partitioning coder and a classified vector quantizer. Note that the set partitioning data is feed into the vector classifier because both of them are determined by the vector energy. For example, if a vector's energy is greater than the current energy threshold (less than the previous energy threshold), a significant bit is sent to indicate the energy range. Based on the energy range, the corresponding energy codebook is selected to code the vector, and the channel symbol is sent to the decoder with the significant bit (i.e., set partitioning data). At the decoder, the same energy codebook is selected based on the received set partitioning data, and the wavelet coefficients are reconstructed through vector de-quantization. After all coefficients are recovered, the decoded image is obtained after inverse wavelet transform.

The key point of the proposed classified vector quantization is that the set partitioning data, which are used to locate the significant vectors, are reused for classifying the vectors. The classification of the vectors is based on the vector energy, and the decoder can select the right codebook to decode the channel symbol. In this way, a significant number of bits are saved for transmitting the side information for vector classification. Since multiple codebooks are available at the decoder, smaller codebooks can be used to reduce the complexity of the vector quantization, while achieving comparable

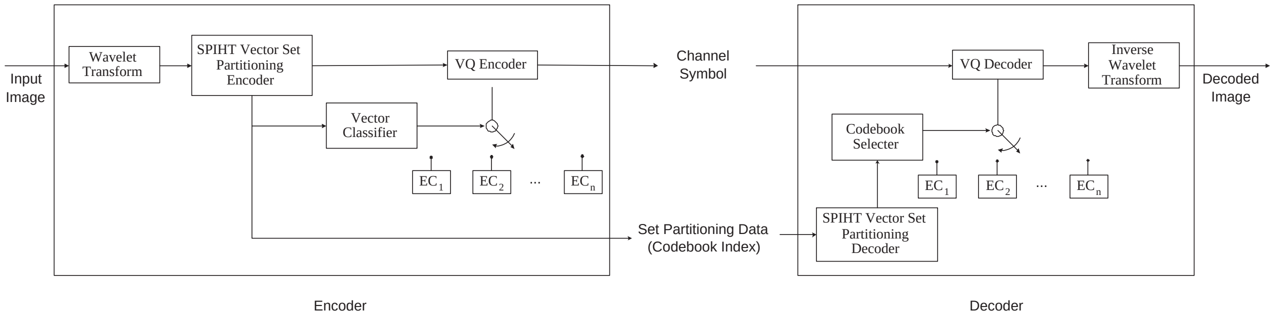

Fig. 2. Encoder and decoder structures of the proposed CVSPIHT.

performance.

Since multiple energy sub-codebooks are used based on the vector energy, there is an important issue of how to choose the size of these codebooks to achieve good image quality. In general, for those vectors with large energies, we intend to use relatively large codebooks to reduce the distortion. However, the size of the codebook also depends on how many training vectors are available. If only a small set of training vectors is available, we can not benefit much from large codebooks. In Section 4, the codebook sizes are given according to the results of an empirical study on a set of test images.

Another issue in our proposed scheme is how to select the energy thresholds to partition the set of vectors. In original SPIHT, since it uses a bit plane coding approach, the threshold to determine whether a coefficient is significant or not has to be a power of 2. After each sorting pass and refinement pass, the threshold is divided by 2. However, in our proposed scheme, since we don't use bit plane coding and no refinement pass is required, the thresholds used to partition the vectors are very flexible. For example, they can be chosen so that each energy training set has roughly the same number of vectors. In this paper, we just use power of 2 thresholds, as shown in Section 4, for simple implementation.

4. EXPERIMENTAL RESULTS

The performance of the proposed CVSPIHT method is demonstrated with natural grayscale images in this section. For comparison, the results obtained by SPIHT and SFQ (implemented by QccPack [11]) are used as a reference. Arithmetic coding is not used in SPIHT and the proposed CVSPIHT methods for a fair comparison.

A 3-level wavelet decomposition implemented by the 7/9 biorthogonal filters [9] is used, which is similar to experiments conducted with SFQ and SPIHT available in [11]. Note that other decomposition levels could also be investigated. Each 2×2 block of wavelet coefficients is grouped together to form a vector of dimension 4. Note that, both codebook generating and encoding processes use the same method to do the wavelet decomposition and vector formation.

To generate the energy codebooks required by the encoding process, a super training set is formed by 2×2 wavelet coefficients from 7 different 512×512 grayscale images, which are Lena, Chair, Elaine, Goldhill, Peppers, Boat, and Anonymous. Then, this super set is split into several energy training sets based on the preset vector energy thresholds, which are shown in Table 1. At last, the energy codebooks are generated by the generalized Lloyd algorithm from the corresponding energy training sets. The codebook sizes are given by Table 1.

Energy range	16-32	32-64	64-128	128-256	256-512	512-1024	1024-∞
Codebook size	32	512	1024	1024	1024	1024	1024

Table 1. The codebook sizes and corresponding energy ranges for different training sets.

Table 2 shows the performance, in terms of PSNR (peak signal to noise ratio), of different image coding methods, including SFQ, SPIHT, and the proposed CVSPIHT, applied to the test images. Note that the images marked with "*" are not in the training set. In terms of PSNR, the proposed CVSPIHT method performs consistently better than SPIHT and SFQ methods for the images both inside and outside of the training set.

Test images	Bit rate	SFQ	SPIHT	CVSPIHT
Lena	0.228	30.24	30.77	33.00
Chair	0.370	29.58	30.16	32.46
Elaine	0.166	29.67	29.88	31.27
Goldhill	0.325	29.61	29.75	31.16
London *	0.379	30.23	30.34	31.55
Peppers	0.233	29.94	30.60	32.46
Boat	0.339	29.80	30.09	31.98
Brian *	0.207	30.68	30.95	31.27
Barbara *	0.463	28.51	28.95	29.12
anonymous	0.222	31.21	30.61	33.46

Table 2. PSNR values (in dB) and bit rates (in bits per pixel) of SFQ, SPIHT, and the proposed CVSPIHT methods.

To examine the subjective image quality, the original Lena image is shown in Fig. 3(a), while the encoded Lena images with SFQ, SPIHT, and the proposed CVSPIHT methods at bit rate of 0.228 bits per pixel are shown in Fig. 3(b)-(d), respectively. It can be seen that the proposed CVSPIHT provides better visual quality than SPIHT and SFQ, such as better edges around the hat and more details of eyes.

The computational complexity of the proposed method depends on the codebook sizes in vector quantization. Since CVQ is used, the codebook sizes are controlled in a desired range. Compared with other components such as DWT and vector SPIHT, it doesn't increase the complexity significantly.

5. CONCLUDING REMARKS

We have presented in this paper a classified vector quantization with SPIHT method for encoding the wavelet coefficients. The reuse of

(a)

(b)

(c)

(d)

Fig. 3. Original and encoded Lena images at bit rate of 0.228 bpp. (a) Original, (b) SFQ (PSNR = 30.24), (c) SPIHT (PSNR = 30.77), (d) CVSPIHT (PSNR = 33.00).

the set partitioning information as the classification index for vector quantization saves a lot of bits. In addition, since we use the vector energy as the state of the encoder and decoder, a perfect match is always found between the sub-codebooks and the state of the encoder/decoder. Experimental results suggest that the proposed method can produce compressed images, inside and outside the training set, with quality better than that generated by other zero tree type methods, such as SFQ and SPIHT.

6. REFERENCES

[1] Z. 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, Dec. 1993.

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

[4] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer, Boston, 1992.

[5] T. Kim, "Side match and overlap match vector quantizers for images," *IEEE Trans. Image Proc.*, vol. 1, no. 2, pp. 170–185, Apr. 1992.

[6] P.C. Cosman, S.M. Perlmutter, and K.O. Perlmutter, "Tree-structured vector quantization with significance map for wavelet image coding," in *Proc. Data Compression Conf.*, Mar. 1995, pp. 33–41.

[7] D. Mukherjee and S.K. Mitra, "Vector spiht for embedded wavelet video and image coding," *IEEE. Trans. Circ. Syst. Video Tech.*, vol. 13, no. 3, pp. 231 – 246, Mar. 2003.

[8] E.Y. Lam, "Statistical modeling of the wavelet coefficients with different bases and decomposition levels," in *IEE Proc. of Vision, Image and Signal Processing*, Jun. 2004, vol. 151, pp. 203 – 206.

[9] P. Mathieu M. Antonini, M. Barlaud and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Proc.*, vol. 1, no. 2, pp. 205 – 220, Apr. 1992.

[10] X. Li J. Knipe and B. Han, "An improved lattice vector quantization scheme for wavelet compression," *IEEE Trans. Signal Proc.*, vol. 46, no. 1, pp. 239 – 243, Jan. 1998.

[11] J.E. Fowler, "Qccpack: an open-source software library for

quantization, compression, and coding," in *Proc. Data Compression Conf.*, Mar. 2000, p. 554.