A NOVEL DATA REPRESENTATION STRATEGY FOR WAVELET IMAGE COMPRESSION

Bing-Bing Chai, Jozsef Vass and Xinhua Zhuang

Department of Computer Engineering & Computer Science
University of Missouri-Columbia, Columbia, MO 65211
{chai,vass,zhuang}@ece.missouri.edu

ABSTRACT

Recent success in wavelet image coding is mainly attributed to recognition of the importance of data organization and representation. Several very competitive wavelet coders have been developed, namely, Shapiro's embedded zerotree wavelets (EZW), Servetto *et al.*'s morphological representation of wavelet data (MRWD), and Said and Pearlman's set partitioning in hierarchical trees (SPIHT). In this paper, we develop a novel wavelet image coder called significance-linked connected component analysis (SLCCA) of wavelet coefficients that exploits both within-subband clustering of significant coefficients and cross-subband dependency in significant fields. Extensive computer experiments show that the proposed SLCCA outperforms all three aforementioned wavelet coders. For example, for the "Barbara" image, at 0.50 bpp SLCCA outperforms EZW and SPIHT by 1.75 dB and 0.89 dB in PSNR, respectively. It is also observed that SLCCA works extremely well for images with large texture regions. For eight typical 256×256 grayscale texture images compressed at 0.40 bpp, SLCCA outperforms SPIHT by 0.32 dB–0.70 dB. This outstanding performance is achieved without any optimal bit allocation procedure. Thus both the encoding and decoding procedures are fast.

1. INTRODUCTION

Conventional wavelet or subband image coders [1, 2] mainly exploit the energy compaction property of subband decomposition by using optimal bit allocation strategies. The drawback is apparent in that all zero-valued wavelet coefficients, which convey little information, must be represented and encoded, biting away a significant portion of the bit budget. Although this type of wavelet coders provide superior visual quality by eliminating the blocking effect in comparison to block-based image coders such as JPEG, their objective performance measured by PSNR increases only moderately.

Two important issues in wavelet coding are:

(a) What is the statistical distribution of a wavelet-transformed image within or across subbands?
(b) How to take advantage of the statistical properties of a wavelet-transformed image?

Empirically, it has been observed that a wavelet-transformed image has the following statistical properties:

1. spatial-frequency localization,
2. energy compaction,
3. within-subband clustering of significant coefficients,
4. cross-subband similarity,
5. decaying of magnitude of wavelet coefficients across subbands.

In recent years, we have seen an impressive advance in wavelet or subband image coding. The success was mainly attributed to the innovative strategies for data organization and representation of wavelet coefficients which exploit not only the energy compaction but also other important statistical properties of wavelet transform. There were three such wavelet image coders published, namely, Shapiro's embedded zerotree wavelet coder (EZW) [3], Servetto *et al.*'s morphological representation of wavelet data (MRWD) [4], and Said and Pearlman's set partitioning in hierarchical trees (SPIHT) [5]. They are all based on empirical observations of the statistical distributions of wavelet-transformed images in one aspect or another. Both EZW and SPIHT exploit cross-subband dependency of insignificant coefficients while MRWD does within subband clustering of significant coefficients. The PSNR of reconstructed images using such data organization strategies was raised by 1–3 dB over block-based transform coders.

In this paper, we propose a novel and more efficient data representation strategy for wavelet image coding termed *significance-linked connected component analysis* (SLCCA). SLCCA strengthens MRWD by exploiting both *within-subband clustering* of significant coefficients and *cross-subband dependency* among significant fields. The cross-subband dependency is effectively exploited by using the so-called significance-link between a parent cluster and a child cluster.

The rest of the paper is organized as follows. Our wavelet image coding algorithm, SLCCA, is presented in next section. In Section 3, the performance of SLCCA is evaluated against three other wavelet coders, i.e., EZW, MRWD, and SPIHT. The last section concludes the paper.

2. SIGNIFICANCE-LINKED CONNECTED COMPONENT ANALYSIS

2.1. Formation of Connected Components within Subbands

Since a rather large portion of wavelet field appears insignificant and significant coefficients within subbands tend to be more clustered (Fig. 2), organizing and representing each subband as irregular shaped clusters of significant coefficients provides an efficient way for encoding. Clusters are progressively constructed by using conditioned dilation, resulting in an effective segmentation of the within-subband significant field. The idea was sketched in [4]. In the following, we will focus our discussion on the selection of structuring elements.

Suppose A is a binary image, B a binary structuring element, and $M \subset A$ a marker. Then, the *conditioned dilation* is defined as

$$D^1(M, A) = (M \oplus B) \cap A,$$

where $\oplus$ denotes the morphological dilation and $\cap$ the intersection. Let

$$D^n(M, A) = D^1(D^{n-1}(M, A), A).$$

Then $D^\infty(M, A)$ defines a cluster in A. For a digital image, the cluster is formed in finite number of iterations when $D^n(M, A) = D^{n-1}(M, A)$.

In the case of clustering in wavelet field, the binary image A represents the significance map, i.e.,

$$A[x, y] = \begin{cases} 1, & \text{if the wavelet coefficient at location} \\ & \text{[x,y] is significant,} \\ 0, & \text{otherwise.} \end{cases}$$

The marker $M \subset A$ represents the seed of a cluster.

Traditionally, a connected component is defined based on one of the three types of connectivity: 4-connected, 8-connected, and 6-connected, each requiring geometric adjacency of two neighboring pixels. Since the significant coefficients in wavelet field are only loosely clustered, the conventional definition of connected component will produce too many components, affecting the coding efficiency. Thus we may use symmetric structuring elements with a size larger than 3×3 square. But we still call the segments generated by conditioned dilation *connected components* even if they are not geometrically connected. Some structuring elements tested in our experiments are shown in Fig. 1. The ones in Figs. 1a and 1b generate 4- and 8-connectivity, respectively. The structuring elements in Figs. 1c and 1d may not preserve geometric connectivity but perform better than the former two in terms of coding efficiency.

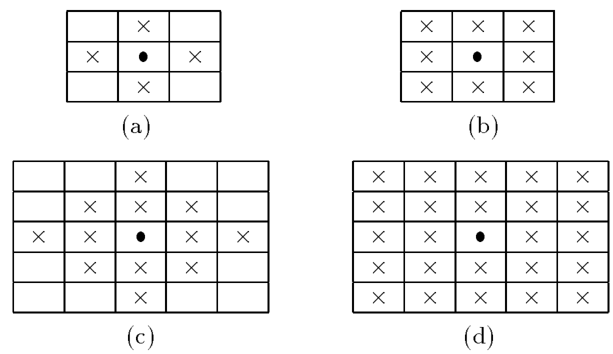
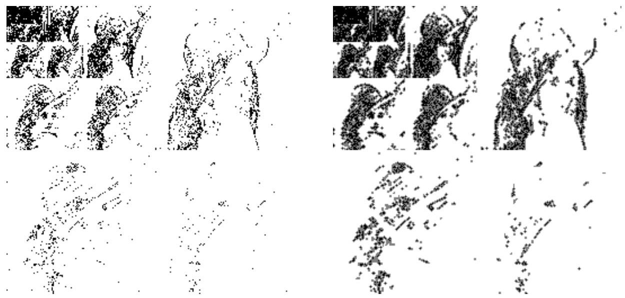


Figure 1: Structuring elements used in conditioned dilation.

To effectively delineate a cluster of significant coefficients, all zero coefficients within the neighborhood B of each significant coefficient in the cluster need to be coded as the boundary of the cluster. By increasing the size of the structuring element, the number of connected components decreases. On the other hand, a larger structuring element results in more boundary zero coefficients. The optimal choice of the size of the structuring element is determined by the cost of encoding boundary zeros versus that of encoding the positional information of connected components. Since the *significance-link* largely reduces the positioning cost, relatively smaller structuring elements can be selected for connected component analysis.

Since extremely small clusters usually do not produce discernible visual effects, and those clusters render a higher insignificant-to-significant coefficient ratio than large clusters, they are eliminated to avoid relatively more expensive coding cost. The connected component analysis is illustrated in Fig. 2. The significance map obtained by quantizing all wavelet coefficients with a uniform scalar quantizer of step size $q = 11$ is shown in Fig. 2a. The 22748 significant wavelet coefficients form 1654 clusters using the structuring element shown in Fig. 1c. After removing connected components having only one significant coefficient, the number of clusters is reduced to 689. The final encoded significance map is shown in Fig. 2b. It is clear that only a small fraction of zero coefficients are encoded.



(a) (b)

Figure 2: Significance map for six-scale wavelet decomposition, $q = 11$. (a) Significance map after quantization: White pixels denote insignificant coefficients and black pixels significant coefficients. (b) The encoded significance map. White pixels denote coefficients that are not encoded. Black and gray pixels denote encoded significant and insignificant wavelet coefficients, respectively.

2.2. Significance-Link in Wavelet Pyramid

The cross-subband similarity among *insignificant coefficients* in wavelet pyramid has been exploited in EZW and SPIHT that greatly improves the coding efficiency. On the other hand, it is found that the spatial similarity in wavelet pyramid is not strictly satisfied, i.e., an insignificant parent does not warrant all four children insignificant. The "isolated zero" symbol used in EZW indicates the failure of such a dependency. The similarity described by zerotree in EZW and the similarity described by insignificant all second generation descendents in SPIHT are more of a reality when a large threshold is used. As was stated in [3] and [6], when the threshold decreases (for embedding) to a certain point, the tree structure or set-partitioned-tree structure are no longer efficient.

In the proposed algorithm, as opposed to EZW and SPIHT, we attempt to exploit the spatial similarity among *significant coefficients*. However, we do not seek a very strong parent-child dependency for each and every significant coefficient. Instead, we try to predict the existence of clusters at finer scales. Statistically, the magnitudes of wavelet coefficients decay from a *parent* to its *children*. It implies that in a cluster formed within a fine subband, there likely exists a significant child whose parent at the coarser subband is also significant. In other words, a significant child can likely be traced back to its parent through this *significance linkage*. It is crucial to note that this significance linkage relies on a much looser spatial similarity.

Formally, two connected components or clusters are called *significance-linked* if the significant parent belongs to one component, and at least one of its children is significant and lies in another component (Fig. 3). If the positional information of the significant parent in the first component is available, the positional information for the second component can be inferred through marking the parent as having a significance-link. Since there are generally many significant coefficients in connected components, the likelihood of finding significance-link between two components is fairly high. Apparently, marking the significance-link costs much less than directly encoding the position, and a significant saving on encoding cluster positions is thus achieved. Among all, using significance-link makes a major difference between SLCCA and MRWD.

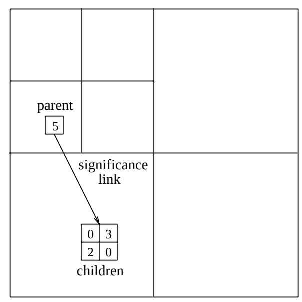


Figure 3: Illustration of significance-link. Nonzero values denote significant coefficients.

2.3. Bit-Plane Organizing and Adaptive Arithmetic Coding

As in most image compression algorithms, the last step of SLCCA involves entropy coding for which adaptive arithmetic coding [7] is employed. Entropy coding techniques attempt to exploit the source statistics in order to generate an average codeword length closer to the source entropy. In contrast to a fixed model arithmetic coder, which works well for a stationary Markov source, the adaptive arithmetic coder updates the conditional probability estimation every time when the coder visits a particular context. For the data stream generated by a nonstationary source such as natural images, the conditional probabilities or local probability distributions may vary substantially from one section to another. The knowledge of the local probability distributions acquired by an adaptive model is more robust than the global estimates and follows the local statistical variation well. In comparison to the fixed model arithmetic coder, the adaptive arithmetic coder is thus able to achieve higher compression. In order to exploit the full strength of an adaptive arithmetic coder, it is preferable to organize outcomes of a nonstationary Markov source into such a stream that each local probability distribution is in favor of one source symbol. This is the basic idea behind the well known lossless bit-plane coding, in which an original image is divided into bit-planes with each bit-plane being encoded separately. Since more significant bit-planes generally contain large uniform areas, the entropy coding techniques can be more efficient.

This idea is employed by SLCCA to encode the magnitude of significant coefficients in each subband. The magnitude of each significant coefficient is converted into a binary representation with a fixed length determined by the maximum magnitude in the subband. Generally, most magnitudes in the subband are smaller than their maximum, implying that more significant bit-planes would contain a lot more 0's than 1's. Accordingly, the adaptive arithmetic coder would generate more accurate local probability distributions in which the conditional probabilities for "0" symbols are closer to one for more significant bit-planes. The context used to define conditional probability models at each significant coefficient is related to the significance status of its eight neighbors and parent.

The bit-plane encoding idea is also used in both EZW and SPIHT but in a different manner. In EZW, for instance, the idea is realized through progressive transmission of magnitudes, with the "0" bits before the first "1" bit being encoded as either "zerotree" or "isolated zero."

3. PERFORMANCE EVALUATION

The SLCCA is evaluated on several natural 512×512 grayscale images. The performance is compared with the best wavelet coders EZW, MRWD, and SPIHT. Each original image is decomposed into a six-scale subband pyramid using the 10/18 filters obtained from ftp.cs.dartmouth.edu. There is no optimal bit allocation carried out in SLCCA. Instead, all wavelet coefficients are quantized with the same uniform scalar quantizer. All the reported bit rates are computed from the actual file sizes.

Table 1 shows the comparison among four wavelet coders at different bit rates. Other results are available at our web site http://www.cecs.missouri.edu/∼dcmms. For "Lena," SLCCA consistently outperforms EZW, MRWD, and SPIHT as well. Compared to EZW, SLCCA gains 1.07 dB in PSNR on average. When compared to MRWD, SLCCA is superior by 0.27–1.07 dB. Compared to SPIHT, SLCCA gains 0.18 dB. For "Barbara," on average, SLCCA is superior to EZW by 1.67 dB, and to SPIHT by 0.62 dB.

Image	Rate [bpp] Algorithm	0.125	0.25	0.50	1.00
Lena	EZW	30.23	33.17	36.28	39.55
	MRWD	-	-	36.60	40.17
	SPIHT	31.09	34.11	37.21	40.41
	SLCCA	31.38	34.33	37.38	40.44
Barbara	EZW	24.03	26.77	30.53	35.14
	SPIHT	24.86	27.58	31.39	36.41
	SLCCA	25.45	28.43	32.28	37.15

Table 1: Performance comparison (PSNR [dB]) of different wavelet coding algorithms.

It appears that SLCCA performs significantly better than SPIHT for images which are rich in texture such as "Barbara." For images which are relatively smooth ("Lena"), the performance between SLCCA and SPIHT gets closer. To further verify the above observation, we compare the performance of SLCCA and SPIHT on eight typical 256×256 grayscale texture images shown in Fig. 4. The results at 0.4 bpp are summarized in Table 2 indicating that SLCCA constantly outperforms SPIHT by 0.32–0.70 dB. An explanation is as follows. When textured images are encoded, wavelet transform is unlikely to yield many large zero regions for lack of homogeneous regions. Thus, the advantage of using an insignificant tree as in EZW, or an insignificant part-of-tree structure as in SPIHT is weakened. On the other hand, SLCCA uses significance-based clustering and significance-based between-cluster linkage, which are not affected by the existence of textures.

Image	SLCCA	SPIHT
"fingerprint"	27.61	27.07
"sweater"	41.83	41.48
"grass"	25.45	24.82
"pig skin"	26.82	26.50
"raffia"	20.93	20.30
"sand"	24.18	23.63
"water"	29.76	29.19
"wool"	26.40	25.70

Table 2: Performance comparison (PSNR [dB]) of SPIHT and SLCCA on texture images at 0.4 bpp.

Finally, we apply SLCCA to fingerprint image compression, which represents a very important issue demanding the best solution. The FBI has developed a fingerprint image compression algorithm called wavelet scalar quantization (WSQ) [8]. At 0.444 bpp or 18:1 compression, SLCCA yields a PSNR of 35.81 dB as opposed to WSQ's 34.43 dB, corresponding to a 1.38 dB

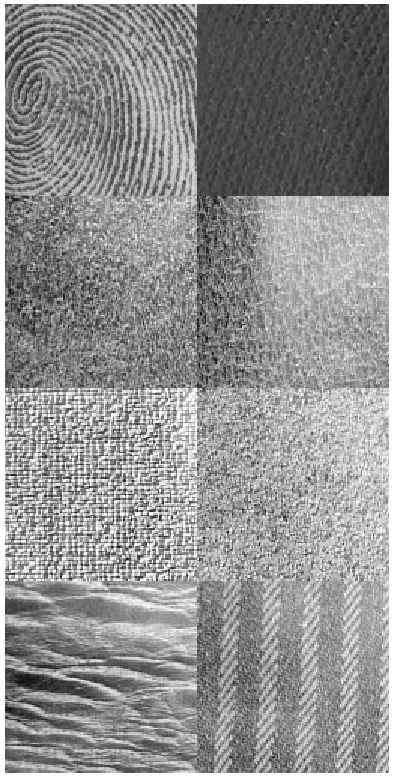
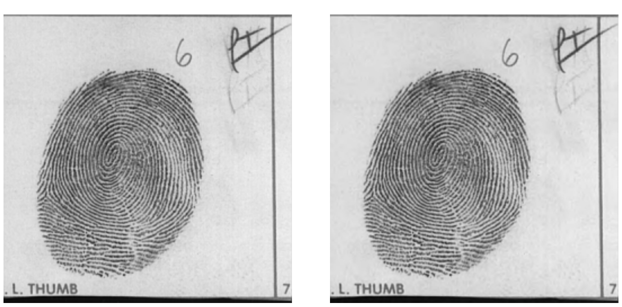


Figure 4: 256×256 texture images. From left to right, top to bottom: "fingerprint," "sweater," "grass," "pig skin," "raffia," "sand," "water" and "wool."

improvement. The original and reconstructed images from SLCCA at 0.444 bpp are shown in Fig. 5. Note that there is almost no loss in visual quality.



(a) (b)

Figure 5: "Fingerprint" images. (a) Original image. (b) Reconstructed image by SLCCA at 0.444 bpp, PSNR=35.57 dB.

4. CONCLUSIONS

A new image coding algorithm termed significance-linked connected component analysis is presented in this paper. The algorithm takes advantage of two properties of the wavelet decomposition: the within-subband clustering of significant coefficients and the cross-subband dependency in significant fields. The significance-link is employed to represent the positional information for clusters at finer scales, which greatly reduces the positional information overhead. The magnitudes of significant coefficients are coded in the bit-plane order so that the local statistic in the bit stream matches the probability model in adaptive arithmetic coding to achieve further saving in bit rate. Extensive computer experiments justify that in most cases, SLCCA surpasses the state-of-the-art image coding algorithms reported in the literature. As no optimization is involved, both the encoding and decoding procedures are remarkably fast.

5. REFERENCES

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

[2] N. Farvardin and N. Tanabe, "Subband image coding using entropy-coded quantization," in *Proceedings of SPIE Image Processing Algorithms and Techniques*, 1990, vol. 1244, pp. 240–254.

[3] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[4] S. Servetto, K. Ramchandran, and M.T. Orchard, "Wavelet based image coding via morphological prediction of significance," in *Proceedings of IEEE International Conference on Image Processing*, Oct. 1995, pp. 530–533.

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

[6] A. Said and W.A. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1303–1310, Sept. 1996.

[7] I.H. Witten, M. Neal, and J.G. Cleary, "Arithmetic coding for data compression," *Communications of ACM*, vol. 30, no. 6, pp. 520–540, June 1987.

[8] J.N. Bradley, C.M. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization standard for grayscale fingerprint image compression," in *Proceedings of SPIE Conference Visual Communications and Image Processing*, 1993.