Highly efficient codec based on significance-linked connected component analysis of wavelet coefficients

Bing-Bing Chai, Jozsef Vass, and Xinhua Zhuang

Department of Computer Engineering & Computer Science
University of Missouri-Columbia
Columbia, MO 65211, USA

ABSTRACT

Recent success in wavelet coding (or interchangeable, linear subband coding) is mainly attributed to the recognition of importance of data organization. There have been several very competitive wavelet codecs 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 propose a new image compression algorithm called Significance-Linked Connected Component Analysis (SLCCA) of wavelet coefficients. SLCCA exploits both within-subband clustering of significant coefficients and cross-subband dependency in significant fields. A so-called significance link between connected components is designed to reduce the positional overhead of MRWD. In addition, the significant coefficients' magnitude are encoded in bit plane order to match the probability model of the adaptive arithmetic coder. Experiments show that SLCCA outperforms both EZW and MRWD, and is tied with SPIHT. Furthermore, it is observed that SLCCA generally has the best performance on images with large portion of texture. When applied to fingerprint image compression, it outperforms FBI's wavelet scalar quantization by about 1 dB.

Keywords: wavelet transform, data organization, spatial similarity, significance-link, connected component, bit plane order

1. INTRODUCTION

It is known that wavelet theory fundamentally identifies subband coding, which has been popularly used for image data compression since the 1980's, as wavelet's discrete cousins. Furthermore, wavelet theory provides an insight into the construction of subband filters, as well as a more productive approach to designing them. The latter is evidenced by the introduction of symmetric biorthogonal wavelet bases with compact support which are instantly converted into more desirable linear phase filters, while still maintaining perfect reconstruction.

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 coefficients in wavelet fields, which convey little information, must be represented and encoded, biting away a significant portion of the bit budget. Although this type of wavelet coder, by the elimination of blocking effect, provides superior visual quality in comparison to block-based image coders, such as JPEG, their objective performance, i.e., PSNR, remains just compatible. A fundamental issue with wavelet coding is: what are the statistical distributions of wavelet-transformed images within or across subbands when images admit a model such as Markov random field? We will address this important issue in a separate paper.

In recent years, we have seen an impressive advance of wavelet or subband image coding. The success is mainly attributed to innovative strategies for data organization and representation of wavelet-transformed images. 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

Further author information-
B-B. C.: E-mail: chai@ece.missouri.edu
J. V.: E-mail: vass@ece.missouri.edu
X. Z. (correspondence): E-mail: zhuang@ece.missouri.edu, Telephone: (573)882-2382, Fax: (573)882-8318

trees (SPIHT).[5] They are all based on empirical observations of the statistical distributions of wavelet-transformed images one way or the other. Both EZW and SPIHT exploit cross-subband dependency of insignificant wavelet coefficients while MRWD exploits within-subband clustering of significant wavelet coefficients. The PSNR of reconstructed images is raised by 1-3 dB over block-based transform coders.

In this paper, we propose a new wavelet image coder called significance-linked connected component analysis (SLCCA) which exploits both within-subband clustering of significant coefficients and cross-subband dependency in the significant field. SLCCA strengthens MRWD by exploiting cross-subband dependency among clusters. The rest of the paper is organized as follows. Section 2 reviews the existing data representation strategies. In Section 3, we present the new image coding algorithm. Section 4 evaluates the performance of SLCCA against other top performance wavelet coders and Section 5 concludes the paper.

2. OVERVIEW OF DATA ORGANIZATION STRATEGIES

There are two distinct types of algorithm for efficient representation of wavelet coefficients reported in the literature. While EZW and SPIHT use tree structured zero regions to represent insignificant coefficients, MRWD deals directly with significant coefficients. A wavelet coefficient c is defined as *significant* if its magnitude is larger than a predefined threshold T, i.e., $|c| \geq T$, otherwise, it is deemed *insignificant*.

EZW efficiently identifies zero regions across subbands while defining the significant field everywhere outside these zero regions by progressive approximation of coefficients' binary representation. Arbitrary shaped zero regions are approximated by the union of highly constrained tree-structured zero regions, each called *zerotree*. Since each *zerotree* can be effectively represented by its root symbol, significant improvement in coding performance is achieved. On the other hand, there may still be many zero coefficients which cannot be included in the highly structured zerotrees. These isolated zeros remain very expensive to represent.

SPIHT seeks to enhance EZW by partitioning the cross-subband tree structure used in EZW into three parts, i.e., tree root, children of the root, and non-child descendants of the root; the last part comprises a majority of the population in the tree. Once a child coefficient is significant, EZW has to represent and encode every non-child descendent separately, even though they may all be insignificant. In SPIHT, as the non-child descendents form a set partitioned from both tree root and children of the root, a single symbol can be employed to represent and encode them as a single unit when they are all insignificant. This fine set partitioning strategy leads to an increase in PSNR by 0.2–0.7 dB over EZW. It can be seen that SPIHT exploits cross-subband dependency more efficiently than EZW.

MRWD is based on the observation that significant coefficients within subbands are more likely to be clustered. In contrast to EZW and SPIHT, MRWD deals directly with the significant field by forming irregular-shaped clusters of significant coefficients in terms of conditioned morphological dilation using a binary structuring element. Although the boundary zeros of each cluster must be represented, the expensive cost of representing and encoding isolated zeros in EZW is avoided. As a result, MRWD constantly outperforms EZW. Nevertheless, MRWD needs to specify a seed for each cluster and encode its positional information as overhead. As a large number of clusters are involved, the overall overhead may take up a significant portion of the bit budget. Among the top three wavelet image coders, SPIHT performs the best in general.

3. SIGNIFICANCE-LINKED CONNECTED COMPONENT ANALYSIS

It has been observed that discrete wavelet transform has the following properties that can be utilized for image compression. All except property 4 can be seen from Fig. 1a.

1. Energy compaction into a small set of wavelet coefficients.
2. Spatial-frequency localization.
3. Clustering of significant coefficients within each subband.
4. Magnitude decay of wavelet coefficients in wavelet pyramid.
5. Cross-subband similarity in wavelet pyramid.

The proposed SLCCA algorithm is an attempt to make use of the above properties.

3.1. Formation of connected components in wavelet field

It has been found empirically that significant coefficients within subbands are more clustered than a 2-D Poisson distribution sharing the same marginal probability,[4] and that statistical clusters can be appropriately constructed by conditioned dilation operation, the latter progressively practicing an efficient segmentation of a significant field. Since the significant coefficients in wavelet field are only loosely clustered together, the conventional definition of connected component will result in too many components, affecting the coding efficiency. Thus we may use symmetric structuring elements with sizes larger than 3×3 square (8-connected), but still call the segments generated by conditioned dilation *connected components* or *clusters*, even though they may not be geometrically connected. Four of the structuring elements tested in our experiments are shown in Fig. 2, where Fig. 2a and Fig. 2b are the 4-neighborhood and 8-neighborhood, respectively, used in the conventional connected component analysis. In order to identify the relative position of significant coefficients within a cluster, all zero coefficients within the neighborhood of each significant coefficient are included in the cluster during encoding.

It is clear that the larger the structuring element, the less the number of connected components in an image. On the other hand, the larger the size of structuring element, the more boundary zero coefficients are included in a component. The optimal choice of the size of structuring element is determined by the cost of encoding boundary zeros versus encoding the positional information of a connected component. In the following, we will present the concept of significance-link that can reduce the cost in encoding the seeds' position. In this way, a relatively smaller structuring element can be selected for connected component analysis.

Clusters with a small number of significant coefficients have a higher insignificant-to-significant coefficient ratio than larger clusters. For example, if the 3×3 structuring element is used, a cluster with only one significant coefficient has eight boundary zeros. Thus small clusters are relatively more expensive to encode than large clusters. In order to avoid the expensive cost of encoding very small clusters, a post-processing is performed where clusters having significant coefficients less than a given number are eliminated by area thresholding.

3.2. Significance-link in wavelet pyramid

Relative to a given wavelet coefficient, all coefficients at coarser (finer) scales of similar orientation corresponding to the same location are called its *ancestors* (*descendents*).[3] The coefficient at the coarse scale is called the *parent* and all coefficients at the next finer scale of similar location are called *children*. It is observed that spatial similarity exists across subbands among insignificant coefficients as well as significant coefficients. Theoretically, by assuming Markov random field as image model, we are able to prove that statistically the magnitudes of wavelet coefficients decay from the parent to its children. It implies that in a significant cluster formed within a finer subband there likely exists a child whose parent belongs to a significant cluster within a coarse subband. That is, a significant *child* can likely be traced back to its *parent* through *significance linkage*. It is crucial to note this significance linkage relies on a much looser spatial similarity. The similarity represented by *zerotree* (or *insignificant family*) in EZW, and similarity represented by both *insignificant family* and *insignificant all second generation descendents* in SPIHT, are more of a reality when a large threshold is used. As stated in Refs. 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.

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, as shown in Fig. 3. If the positional information of the significant parent in the first component is available, then the positional information for the second component can be inferred when the parent is marked as having a significance-link. Apparently, marking costs much less than encoding the position and saving on positioning clusters is then achieved. Furthermore, as there are generally many significant coefficients in a component, the likelihood of finding significance-link between two components is high. The saving increases as the bit rate increases, ranging from 527 bytes (at 0.25 bpp) to 1183 bytes (at 0.5 bpp) for "Lena" in comparison to MRWD.

3.3. Bit plane ordered representation of wavelet coefficients

In general, the last step of an image coding algorithm involves entropy coding, such as Huffman coding,[7] or arithmetic coding.[8] The zero-order *entropy* $H(S)$ for a given source $S = \{s_1, s_2, \cdots, s_n\}$ is given by

$$H(S) = -\sum_{i=1}^{n} p(s_i) \log_2 p(s_i),$$

where $p(s_i)$ is the probability that symbol s_i occurs in the source stream. Entropy can also be interpreted as the average cost to encode a symbol in the source stream. The underlying idea of entropy coding is to encode rarely appearing symbols with long codewords and frequently occurring symbols with short codewords. Since in practice, the probabilities of the symbols in source alphabet are different, entropy coding is beneficial.

In the proposed algorithm, an adaptive arithmetic coder is employed for final entropy coding. The adaptive arithmetic coder updates its probability model based on the coding history. Adaptive arithmetic coding is experimentally proved to outperform the fixed model arithmetic coding. If the symbols in the source stream can be ordered in such a way that the local probability distribution is heavily in favor of one symbol, the total cost of encoding the source stream can be reduced. This is the idea for bit plane encoding of the magnitudes of significant coefficients described below.

A common way of encoding coefficient magnitudes is to convert each coefficient into a binary representation with a predetermined length and encode the binary string corresponding to each coefficient sequentially. In reality, most of the coefficients are smaller than the maximum value that can be represented by the fixed-length binary string, meaning "0" symbols have to be added to the binary string preceding the first "1" symbol. Thus the binary representation of the coefficients has more 0's than 1's in most significant bits. This fact can be explored by the idea of bit plane encoding. Consider an image or subimage in which each pixel is represented by n bits. By selecting a single bit from the same position in the binary representation of each pixel value, say the kth bit $(0 \leq k < n)$, a binary image called a *bit plane* is formed. If ordering the magnitude of the coefficients in the bit plane order, the bit planes corresponding to the most significant bits will have a higher probability of 0's than 1's when encoding the most significant bit planes. Thus the adaptive arithmetic coding will have a probability model heavily in favor of the "0" symbols for the most significant bit planes. By taking advantage of the local statistics, a saving in bit rate is achieved.

Actually, bit plane encoding is used in both EZW and SPIHT. In both algorithms, bit plane encoding is realized by progressive transmission of magnitudes, where the "0" bits before the first "1" bit are not encoded explicitly. In the proposed algorithm, even though encoding the 0's at the most significant bits is inevitable, bit plane ordering makes it very efficient to encode these 0's.

3.4. SLCCA algorithm

Now we present a complete algorithm for the proposed significance-linked connected component representation scheme.

Step 1. Form a subband pyramid of six scales.

Step 2. Conduct connected component analysis to form clusters of significant coefficients in the subband pyramid. Each cluster is confined within a single subband. Remove the connected components with areas smaller than a predefined threshold A.

Step 3. Starting from the coarsest subband (the top of the pyramid), scan the subbands in the order LL, LH, HL, HH (Fig. 1b) for significant coefficients. Scan the coefficients in the same subband in the top to bottom, left to right order. Go to the next finer scale of subbands after all the coefficients in the current scale have all been scanned and processed.

Step 4. If a coefficient c is found to be significant and has not been encoded, go to Step 5; or else, continue the scan as described in Step 3 and repeat Step 4.

Step 5. Encode the position $[x, y]$ of c.

Step 6. Encode the sign (POS,NEG) of c.

Step 7. If c is the parent of a child cluster that has not been assigned to any other coefficients,

encode a special symbol (LINK);

assign the child cluster to c;

store the child position in a first-in-first-out (FIFO) queue.

Step 8. Conditioned dilation.

For every $[\Delta x, \Delta y]$ in a predefined neighborhood, do

If $[x + \Delta x][y + \Delta y]$ is significant and has not been encoded,

then $x = x + \Delta x$, $y = y + \Delta y$, and go to Step 6.

If $[x + \Delta x][y + \Delta y]$ is insignificant,

then encode a ZERO symbol.

Step 9. If the FIFO queue is not empty, take the next cluster out of the queue, and go to Step 6; or else continue the scan for the next significant coefficient and go to Step 4.

Step 10. Encode the magnitudes of each significant coefficient with the adaptive arithmetic coder.

4. PERFORMANCE EVALUATION

The performance of SLCCA is evaluated on two 512×512 gray scale images "Lena" and "Barbara" (Fig. 4) and compared to the best wavelet coders, EZW, MRWD, and SPIHT. In experiments, the original image is decomposed into a 6-scale subband pyramid using the 9-tap spline variant filters with less dissimilar length provided in Ref. 1. Wavelet coefficients are quantized with a uniform scalar quantizer.

Table 1 shows the PSNR comparison among four wavelet coders on the "Lena" image. SLCCA consistently outperforms both EZW and MRWD. The PSNR is 0.55 – 0.7 dB higher than EZW and at least 0.4 dB higher than MRWD. On the other hand, SPIHT is superior to SLCCA by about 0.16 dB. Table 2 shows the PSNR comparison between SLCCA and SPIHT on the "Barbara" image. It is seen that SLCCA outperforms SPIHT by 0.1 – 0.4 dB. Since "Barbara" has a lot of texture regions while "Lena" is relatively smooth, the experiments seem to imply that SLCCA would perform better when images have a large portion of texture. When textured images are encoded, wavelet transform is unlikely to yield many large zero regions due to the lack of homogeneous regions. Thus, the advantage of using insignificant tree or insignificant part-of-tree structure 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.

Finally, we applied SLCCA to fingerprint image compression. Fingerprint image compression is a very important issue, as the digitized fingerprints of one person may requires 10 MB of storage without compression. The FBI has developed a fingerprint image compression algorithm called wavelet scalar quantization (WSQ)[9] that can compress the image by 18:1 (0.44 bpp) with a PSNR of 34.43 dB (http://www.c3.lanl.gov/~brislawn/FBI/OverCompress/HowItWorks/howitworks.html). As a comparison, SLCCA yields a PSNR of 35.39 dB at the same bit rate, which is 0.96 dB higher than WSQ. The original image and reconstructed image from SLCCA are shown in Fig. 5. Note that there is very little loss in visual quality.

5. CONCLUSIONS

A new image coding algorithm termed SLCCA is proposed in this paper. The algorithm takes advantage of two properties of a wavelet pyramid: the within-subband clustering of significant coefficients and the cross-subband dependency in significance fields. The significance-link is employed to represent the positional information for clusters at finer subbands, which greatly reduces the positional information overhead. The magnitudes of significant coefficients are coded in the bit plane order so that the local statistics in the bit stream match the probability model in adaptive arithmetic coding to achieve further saving in bit rate. Comparison of SLCCA with other state-of-the-art wavelet coding algorithms reveals that SLCCA has better or comparable performance. Our experiment on fingerprints indicates that SLCCA can be a good choice for fingerprint compression.

Our future research direction is to search for a better way of representing the clusters in wavelet fields in order to reduce the cost of encoding the clusters' boundary zeros. SLCCA has to encode the zero boundary of each cluster, which still takes up a significant portion of the bit budget. A possible improvement may be using chain code to represent the cluster boundary and merging clusters close to each other.

REFERENCES

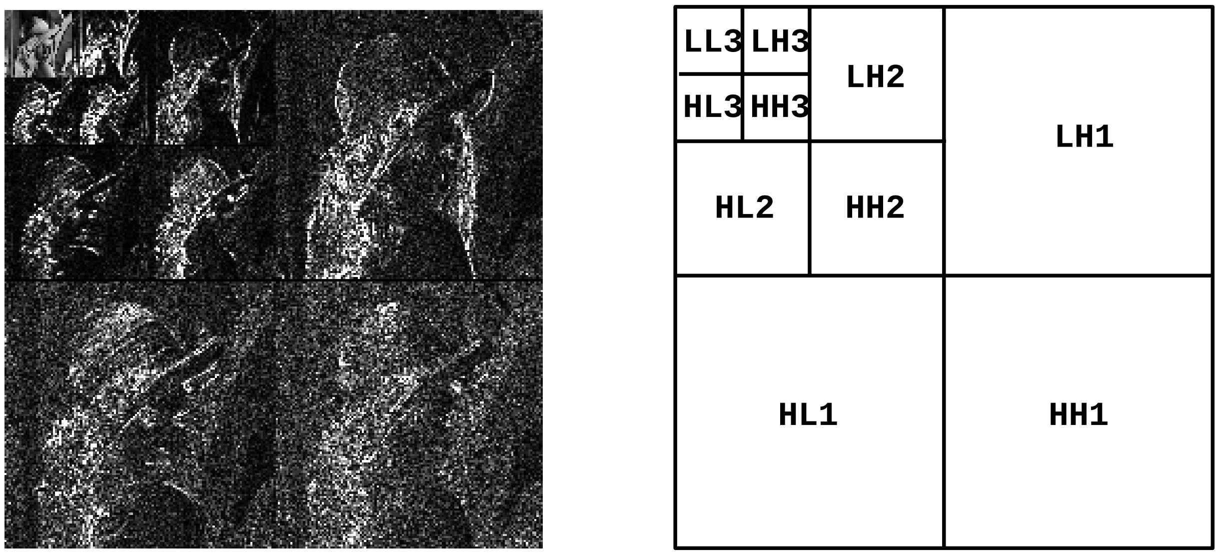
1. M. Antonini, M. Barlaud, P. Mathien, and I. Daubechies., "Image coding using wavelet transform," *IEEE Trans. Image Processing*, Vol. 1, No. 2, pp. 205–221, Apr. 1992.
2. N. Farvardin and N. Tanabe, "Subband image coding using entropy-coded quantization," *SPIE Image Proc. Alg. Tech.*, Vol. 1244, pp. 240-254, 1990.
3. J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, Vol. 41, No. 12, pp. 3445- 3462, Dec. 1993.
4. S. Servetto, K. Ramchandran, and M. Orchard, "Wavelet based image coding via morphological prediction of significance," *Proceeding ICIP-95*, pp. 530-533, Washington, D.C., Oct. 23-26, 1995.
5. A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circ. Sys. Video Techn.*, 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 Trans. Image Processing*, Vol. 5, No. 9, pp. 1303-1310, Sept. 1996.
7. D. A. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings IRE*, Vol. 40, pp. 1098-1101, 1952.
8. I. H. Witten, R. M. Neal, and J. Cleary, "Arithmetic coding for data compression," *Commun. ACM*, Vol. 30, No. 6, 1987.
9. J. N. Bradley, C. M. Brislawn and T. Hopper, "The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression," *Proc. SPIE Conf. on Visual Commun. Image Proceeding*, Orlando, FL, Apr. 1993.

Table 1. Performance comparison (PSNR [dB]) for "Lena" image.

Algorithm	0.25 bpp	0.30 bpp	0.40 bpp	0.50 bpp
EZW[3]	33.17			36.28
MRWD[4]		34.07		36.60
SPIHT[5]	34.11	34.95	36.24	37.21
SLCCA	34.00	34.79	36.09	37.09

Table 2. Performance comparison (PSNR [dB]) for "Barbara" image.

Algorithm	0.25 bpp	0.30 bpp	0.40 bpp	0.50 bpp
SPIHT[5]	28.13	29.18	30.76	32.10
SLCCA	28.56	29.42	30.92	32.24



(a) (b)

Figure 1. Wavelet pyramid. (a)Three-scale wavelet decomposition for "Lena" image. Light color: large magnitude, dark color: small magnitude. (b) Illustration of subbands at different scales.

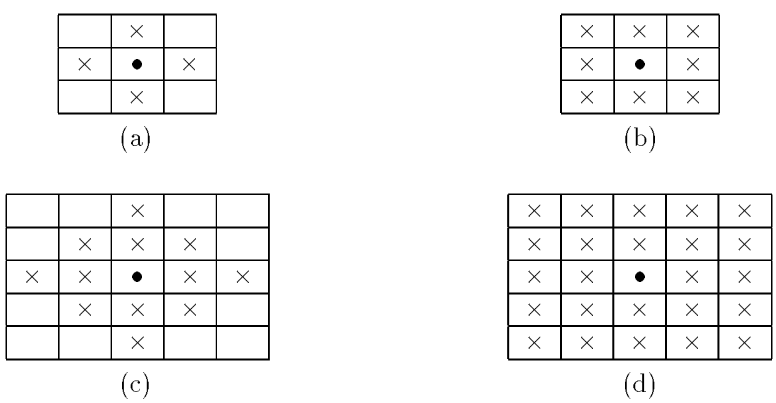


Figure 2. Structuring elements used in conditional dilation.

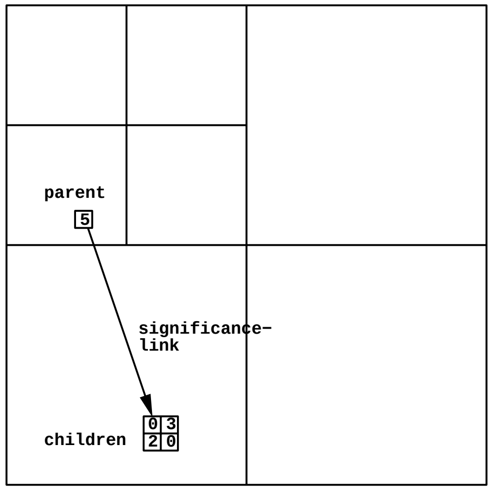


Figure 3. Significance-link. The values are the magnitudes of quantized coefficients. Nonzero values denote significant coefficients.

(a)

(b)

Figure 4. Test images. (a) Lena. (b) Barbara.

(a) (b)

Figure 5. Coding result of fingerprint image. (a) Original image. (b) Reconstructed image of SLCCA (0.44bpp).