

Statistically Adaptive Wavelet Image Coding

Bing-Bing Chai *Jozsef Vass* *Xinhua Zhuang*
Department of Computer Engineering & Computer Science
University of Missouri-Columbia, Columbia, MO 65211

1 Introduction

In the mid-1980's, wavelet theory was developed in applied mathematics [1, 2, 3]. Soon, subband coding [4], which has been a very active research area for image and video compression, was identified as wavelet's discrete cousin. Furthermore, a fundamental insight into the structure of subband filters was developed from wavelet theory that led to a more productive approach to designing the filters [1, 5, 6]. Thus subband and wavelet are often used interchangeably in the literature.

Two types of subband decomposition are commonly used in image compression, i.e., uniform and pyramidal decomposition. Uniform decomposition [7] divides an image into equal-sized subbands (Fig. 1a). By contrast, pyramidal decomposition represents an octave-band (dyadic) decomposition, offering a multiresolution representation of the image as illustrated in Fig. 1b. Most of the subband image coders published recently are based on pyramidal decomposition.

Conventional wavelet or subband image coders [5, 8] 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 [9], their objective performance measured by peak signal-to-noise ratio (PSNR, eq. (1)) increases only moderately.

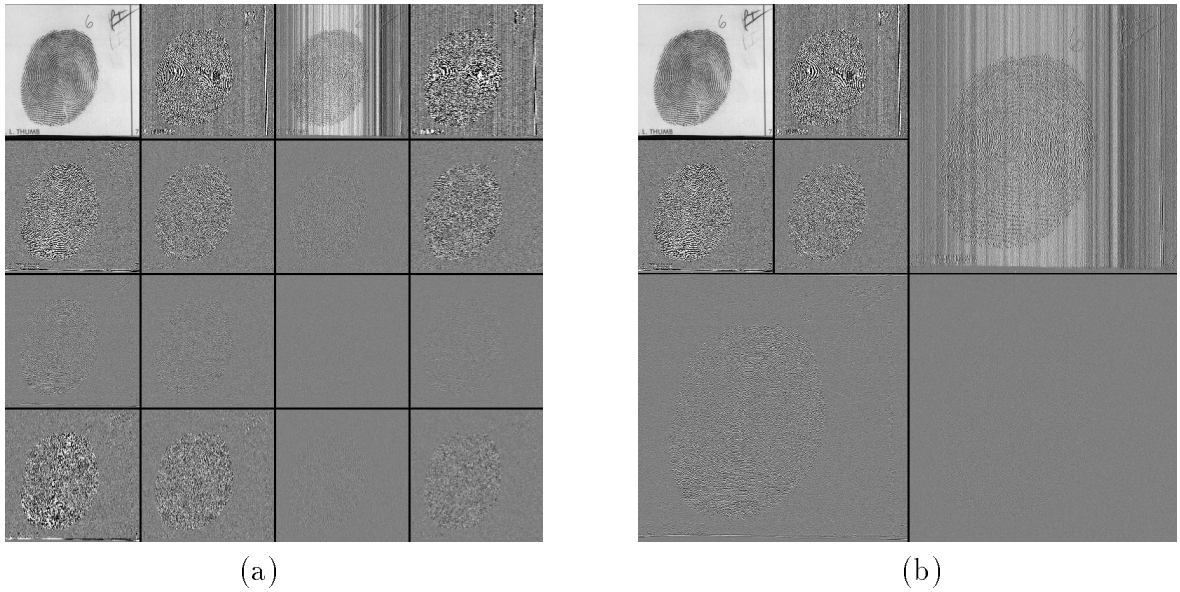


Figure 1: (a) Uniform wavelet decomposition. (b) Pyramidal wavelet decomposition.

A fundamental issue in wavelet coding is: what is the statistical distribution of a wavelet-transformed image within or across subbands when the image admits a stochastic model such as Markov random field? We will brief our initial exploration in this paper. 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 coefficients magnitudes across subbands.

In recent years, we have seen an impressive advance in wavelet image coding. The success is mainly attributed to innovative strategies for data organization and representation of wavelet-transformed images which exploit the above statistical properties one way or the other. Three such top-ranked wavelet image coders have been published, namely, Shapiro's embedded zerotree wavelet coder (EZW) [10], Servetto *et al.*'s morphological

representation of wavelet data (MRWD) [11], and Said and Pearlman’s set partitioning in hierarchical trees (SPIHT) [12]. Both EZW and SPIHT exploit cross-subband dependency of *insignificant* wavelet coefficients while MRWD does within-subband clustering of *significant* wavelet coefficients. As a result, the PSNR of reconstructed images is consistently raised by 1–3 dB over block-based transform coders.

In this paper, we present a novel strategy for data organization and representation for wavelet image coding termed significance-linked connected component analysis (SLCCA). SLCCA adaptively exploits the statistical properties of wavelet-transformed images at each stage of the coding process. SLCCA strengthens MRWD by exploiting not only within-subband clustering of significant coefficients, but also cross-subband dependency in the 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 key components of SLCCA include multiresolution discrete wavelet image decomposition, connected component analysis within subbands, significance-link registration across subbands, and bit-plane encoding of magnitudes of significant coefficients by adaptive arithmetic coding.

The rest of the paper is organized as follows. In Section 2, the aforementioned statistical properties of wavelet-transformed images are first discussed. Then, the data organization and representation strategies used in EZW, MRWD, and SPIHT are analyzed. Our wavelet image coding algorithm, SLCCA, is presented in Section 3. In Section 4, the performance of SLCCA is evaluated against other wavelet coders. The last section concludes the paper.

2 Statistical Properties of Wavelet-Transformed Image and Their Application in Image Compression

The wavelet transform cuts up signals into different frequency bands and then analyzes each band with a resolution matched to its scale. As an adaptive alternative to the classical Short-Time Fourier Transform (STFT) [13], wavelet transform is of primary

interest for the analysis of non-stationary signals. In contrast to STFT, which uses a single analysis window, wavelet transform provides constant relative bandwidth frequency analysis, namely, it uses short time windows at high frequencies and long time windows at low frequencies. When wavelet transform is applied to images, the spatial domain naturally replaces the time domain.

2.1 Statistical Properties of Wavelet-Transformed Images

Discrete-wavelet-transformed images demonstrate the following statistical properties and their exploitation continually proves to be important for image compression.

2.1.1 Spatial-Frequency Localization

Although both FIR and IIR filters can be used for subband decomposition, FIR subband filters are most commonly used. Thus our discussion here will be based on FIR filters. When FIR filters are used, each wavelet coefficient contains only features from a local segment of an input image. Since subband coding decomposes an image into a few frequency bands with almost no overlap, each subband is frequency localized with nearly independent frequency content. In brief, each wavelet coefficient represents information in a certain frequency range at a certain spatial location.

2.1.2 Energy Compaction

A natural image is typically composed of a large portion of homogeneous and textured regions, and a rather small portion of edges including perceptually important object boundaries. Homogeneous regions have the least variation and mostly consist of low frequency components; textured regions have moderate variation and consist of a mixture of low and high frequency components; and edges show the most variation and are mainly composed of high frequency components. Accordingly, wavelet transform compacts most energy distributed over homogeneous and textured regions into the lowpass subband. Each time a lowpass subband at a fine resolution is decomposed into four subbands at a coarser resolution, critical sampling is applied that allows the newly generated lowpass

subband to be represented by only one fourth of the size of the original lowpass subband. By applying this decomposition process repeatedly for a few times on an image, the energy will be effectively compacted into few wavelet coefficients. Fig. 2 shows the one- and two-scale wavelet decompositions of “Lena”’s feather as well as the corresponding energy distribution surfaces. After the two-scale decomposition, most energy is well compacted into one sixteenth of the total wavelet coefficients.

2.1.3 Within-Subband Clustering of Significant Coefficients

A wavelet coefficient c is called *significant* with respect to a predefined threshold T if $|c| \geq T$; otherwise, it is deemed *insignificant*. An insignificant coefficient is also known as zero coefficient. Empirically, it has been found that significant coefficients within subbands are more clustered than a 2-D Poisson distribution which shares the same marginal probability [11]. Due to the absence of high frequency components in homogeneous regions and the presence of high frequency components in textured regions and around edges, significant coefficients in highpass subbands usually appear at the spatial locations of edges or textures of high energy. In other words, they are indicative of prominent “discontinuity” or prominent “changes”, a phenomenon which tends to be clustered. The within-subband clustering of “Lena” image is shown in Fig. 3a. Note, how the clusters of significant coefficients in highpass subbands delineate the contour of “Lena.”

2.1.4 Cross-Subband Similarity

According to the definitions in [14, 10], relative to a given wavelet coefficient, all coefficients at finer scales which correspond to the same spatial location are called its *descendants*; accordingly, the given coefficient is called their *ancestor*. Specifically, the coefficient at the coarse scale is called the *parent* and all four coefficients corresponding to the same spatial location at the next finer scale of similar orientation are called *children* (Fig. 4). Although the linear correlation between the values of parent and child wavelet coefficients has been empirically found to be extremely small as expected, there is likely additional dependency between the magnitudes of parent and children. Experi-

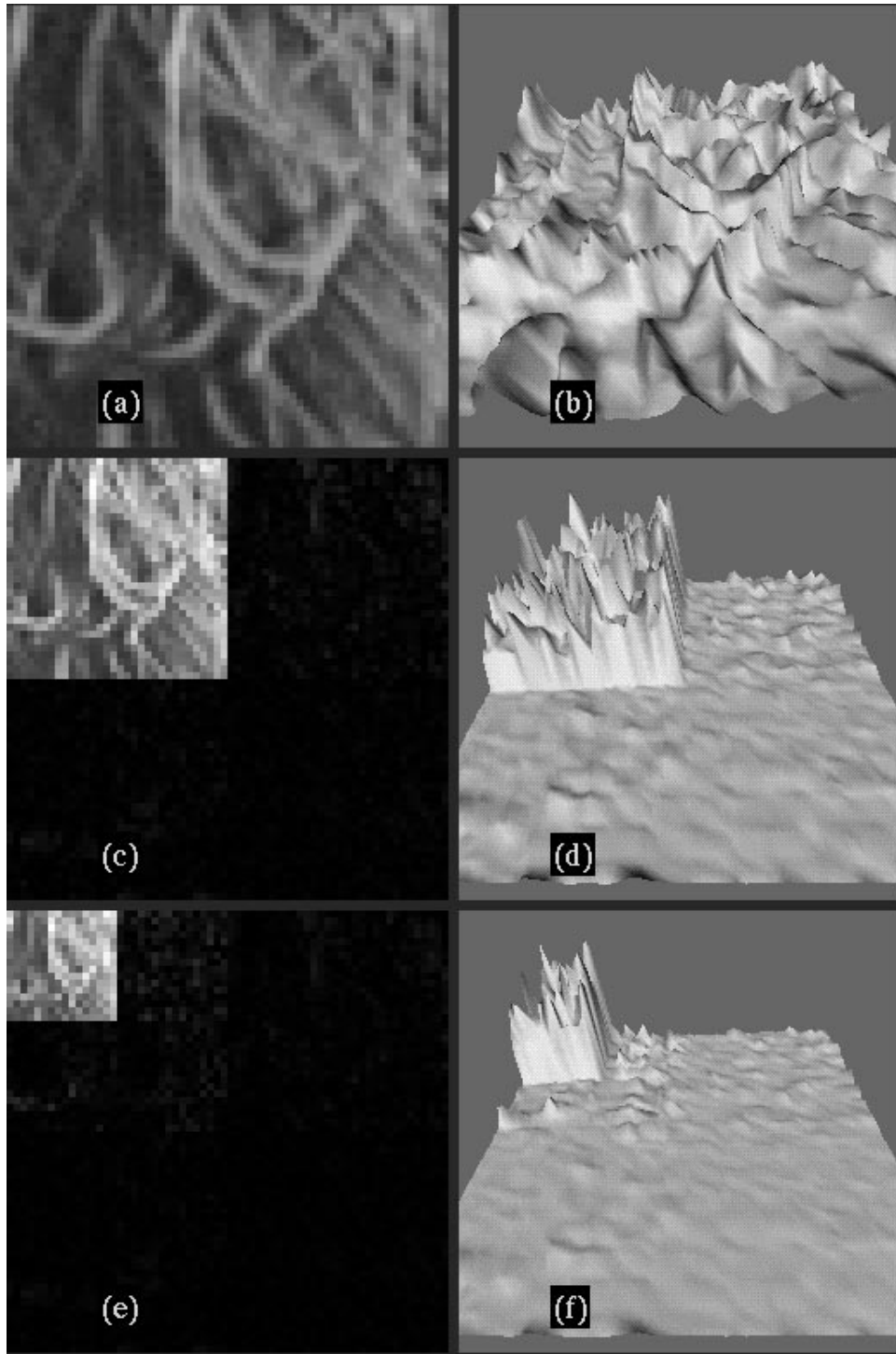


Figure 2: Energy compaction property of wavelet decomposition. (a) Part of the “Lena” image and (b) its surface plot. (c) One-scale and (e) two-scale wavelet decomposition and the corresponding surface plots are shown in (d) and (f), respectively.

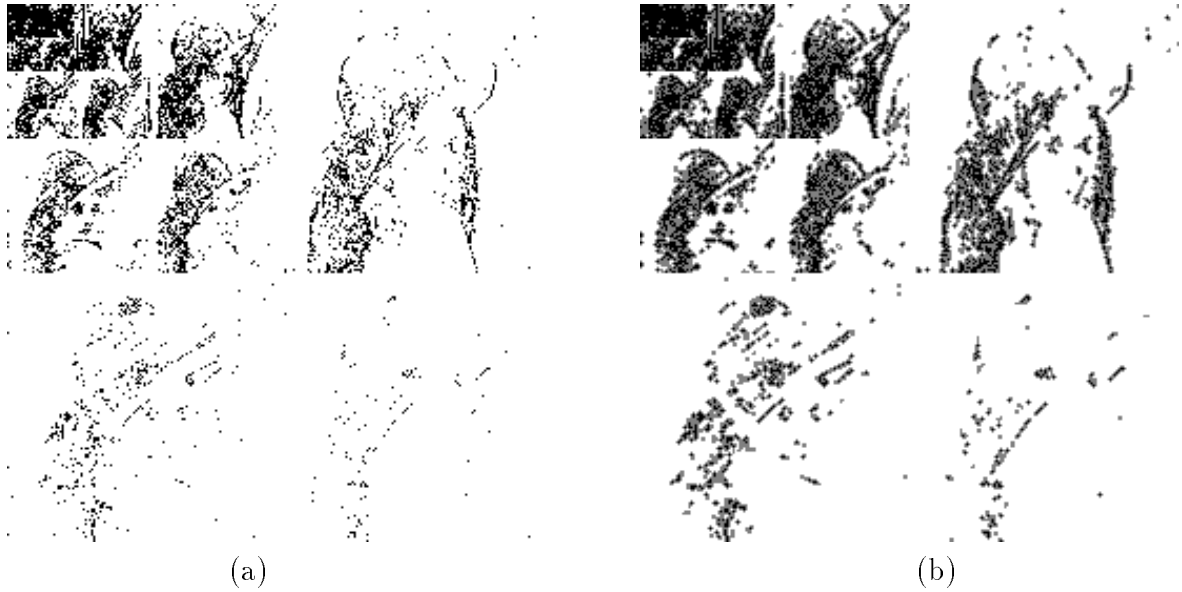


Figure 3: Significance map for six-scale wavelet decomposition, $q = 11$. (a) Significance map after quantization: White pixels denote insignificant coefficients and black pixels denote significant coefficients. (b) The transmitted significance map (after removing clusters having only one significant coefficient). White pixels denote insignificant coefficients that are not encoded. Black and gray pixels denote encoded significant and insignificant wavelet coefficients, respectively.

ments showed that the correlation coefficient between the squared magnitude of a child and the squared magnitude of its parent tends to be between 0.2 and 0.6 with a strong concentration around 0.35 [10].

2.1.5 Decaying of Coefficients Magnitudes Across Subbands

Although it appears to be difficult to characterize and make a full use of the cross-subband magnitude similarity, a reasonable conjecture based on experience with real-world images is that the magnitude of a child is generally smaller than the magnitude of its parent. By assuming Markov random field as the image model, we are able to prove that statistically the magnitude of wavelet coefficients exponentially decays from the parent to its children [15, 16], providing a strong theoretical support to EZW, SPIHT, and SLCCA.

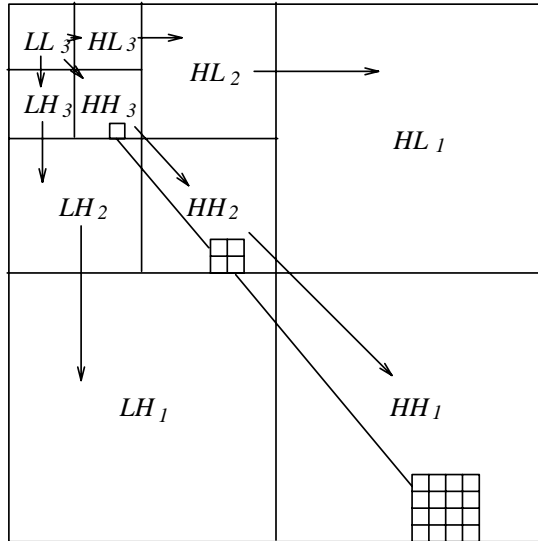


Figure 4: Illustration of parent-child relationship between subbands at different scales. The pixel drawn in HH_3 is the parent of the four pixels in HH_2 .

2.2 Overview of Data Organization and Representation Strategies

Although all wavelet-based image coding algorithms take advantage of energy compaction and spatial-frequency localization properties of wavelet-transformed images, the greatest contribution to improvement in coding efficiency is the exploitation of the within-subband and cross-subband statistical properties. There exist two efficient approaches to the organization and representation of wavelet coefficients in the literature. While EZW and SPIHT use the regular tree structure to approximate the spatial similarity in insignificant fields across subbands, MRWD finds irregular clusters of significant fields within subbands. Among the top three wavelet image coders, SPIHT performs the best in general.

It is commonly accepted from the source coding theory that in general, an image compression technique grows computationally more complex as it becomes more efficient. EZW interrupts this tendency by achieving outstanding performance with very low computational complexity. It efficiently identifies and approximates arbitrary shaped regions with wavelet coefficients equal to zero, i.e, zero regions, across subbands by the

union of highly constrained tree-structured zero regions called *zerotrees* (the structure formed by the pixels in HH_3 , HH_2 , and HH_1 in Fig. 4). Meanwhile, it defines the significant fields outside these zero regions through progressively refining the magnitudes of coefficients. It is apparent that each *zerotree* can be effectively represented by its root symbol. On the other hand, there may still be many zero coefficients which cannot be included in the highly structured zerotrees. These isolated zeros remain expensive to represent.

SPIHT seeks to enhance EZW by partitioning the cross-subband tree structure into three parts, i.e., tree root, children of the root, and non-child descendents of the root, which are illustrated by the pixels shown in HH_3 , HH_2 , and HH_1 in Fig. 4, respectively. It is obvious that the non-child descendents comprise a majority of the population in the tree structure. When a child coefficient is found significant, EZW represents and encodes all four grandchild coefficients separately even if all non-child descendents are insignificant. By contrast, SPIHT treats the insignificant non-child descendents as a union and employs a single symbol to represent and encode it. This fine set partitioning strategy leads to an impressive increase in PSNR by 0.86–1.11 dB over EZW on “Lena” image (see Table 4), indicating that SPIHT exploits cross-subband dependency more efficiently than EZW.

Different from EZW and SPIHT, MRWD directly forms irregular-shaped clusters of significant coefficients within subbands. The clusters within a subband are progressively delineated by insignificant boundary zeros through morphological conditioned dilation operation, which utilizes a structuring element to control the shape and size of clusters, as well as the formation of boundaries. For most structuring elements, a formed cluster could be neither 4-connected nor 8-connected. With MRWD, the boundary zeros of each cluster still need to be coded but the expensive cost of representing and encoding isolated zeros in EZW is largely avoided. As a result, MRWD constantly outperforms EZW. For instance, it gains 0.32–0.62 dB over EZW on “Lena” image, as shown in Table 4. Nevertheless, MRWD *does* need to specify a seed (a pixel from which a cluster is

originated) for each cluster and encode its positioning information as overhead. As a large number of clusters are involved, the overall overhead may take up a significant portion of the bit budget.

3 Significance-Linked Connected Component Analysis

SLCCA attempts to exploit all the aforementioned statistical properties of wavelet-transformed images. In this section, the key features of our wavelet coder SLCCA are first described. Then a complete algorithm is presented.

3.1 Formation of Connected Components within Subbands

It has been seen, that a rather large portion of wavelet coefficients are usually insignificant; and significant coefficients within subbands are more clustered than points from a 2-D Poisson distribution having the same marginal probability [11] (Fig. 3a). Therefore, 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 [17], resulting in an effective segmentation of the within-subband significant regions. This idea was sketched in [11]. In the following, we discuss the issue in regard to the selection of structuring elements.

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

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

where \oplus denotes the morphological dilation [18, 19] 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 iterations when $D^n(M, A) = D^{n-1}(M, A)$.

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

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

The marker $M \subset A$ represents the seeds of each 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 a 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. 5, the ones in Figs. 5a and 5b generate 4- and 8-connectivity, respectively. The structuring elements in Figs. 5c and 5d represent a diamond of size 13 and a 5×5 square, respectively. The latter two may not preserve geometric connectivity within a component, but perform better than the former in terms of coding efficiency.

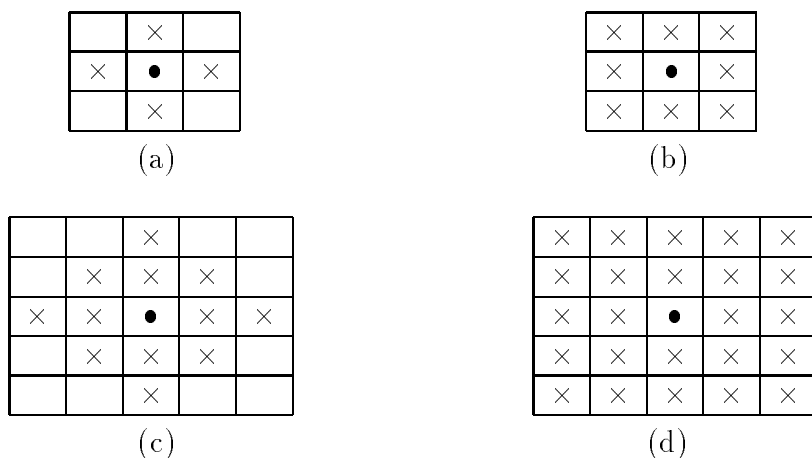


Figure 5: Structuring elements used in conditioned dilation.

To effectively delineate a significant cluster, all zero coefficients within the neighborhood B of each significant coefficient in the cluster need to be marked 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 technique is going to reduce the positioning cost a lot, relatively smaller structuring elements can be selected for connected component analysis.

Area Threshold	Number of Clusters	Significant Coefficients	Transmitted Insignificant Coefficients	Performance PSNR [dB]
0	452	11200	38690	34.26
1	253	11545	36290	34.35
2	188	11638	35028	34.34
3	152	11703	34301	34.32

Table 1: Performance comparison of different area thresholds on Lena image at 0.25 bpp, using 5×5 structuring element (Fig. 5d).

As 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 more expensive coding cost. As the area threshold increases, the number of clusters decreases which results in the reduction of the required cluster positioning information. As illustrated in Table 1 zero area threshold has the worst performance. All other area thresholds have similar performance.

In the paper by Luo *et al.* [20], the authors also propose wavelet coefficients clustering for image compression. They use clustering as a tool for quantization, i.e., wavelet coefficients are clustered together and quantized to the mean value of the given cluster. The wavelet coefficients are then coded by either using traditional runlength coding or Shapiro’s EZW algorithm. In SLCCA, we use clustering to register and transmit the significance map, i.e., clustering is our tool for data organization. The two algorithms

justify each other in the sense that clusters having perceptually less significant information can be removed to raise the coding gain without compromising the objective and subjective quality.

The connected component analysis is illustrated in Fig. 3. The significance map obtained by quantizing all wavelet coefficients with a uniform scalar quantizer with step size $q = 11$ is shown in Fig. 3a. The 22748 significant wavelet coefficients form 1654 clusters using the structuring element shown in Fig. 5c. 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. 3b. It is clear that only a small fraction of zero coefficients are encoded.

3.2 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 both zerotree and insignificant all non-child descendents in SPIHT are more of a reality when a large threshold is used. As was stated in [10] and [21], 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. As pointed out before, 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.

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. 6). 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. The saving from using significance-link increases as the bit rate increases, ranging from 527 bytes (at 0.25 bpp) to 3103 bytes (at 1 bpp) for “Lena” image. Among all, using significance-link makes a major difference between SLCCA and MRWD.

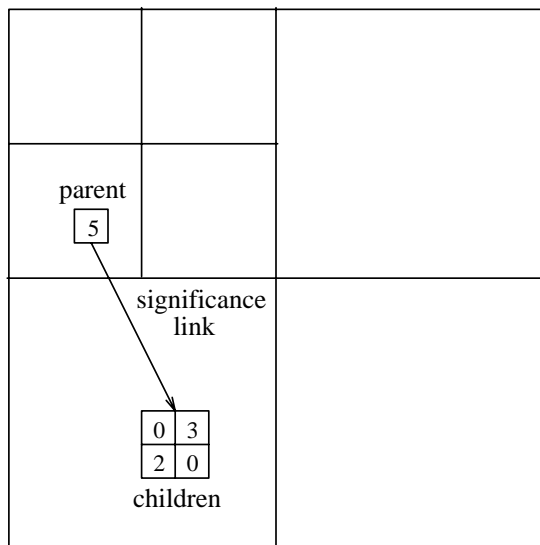


Figure 6: Illustration of significance-link. The values are the magnitudes of quantized coefficients. Nonzero values denote significant coefficients.

3.3 Bit-Plane Organization and Adaptive Arithmetic Coding

As in most image compression algorithms, the last step of SLCCA involves entropy coding for which adaptive arithmetic coding [22] is employed. Adaptive arithmetic coder updates the corresponding conditional probability estimation every time when the coder visits a particular context. Thus the local probability distributions is well exploited, resulting in higher compression than that achieved by fixed model arithmetic coder.

In order to exploit the full strength of an adaptive arithmetic coder, it is preferable to organize outcomes of a non-stationary Markov source into such a stream that local probability distribution is in favor of one source symbol in a certain segment of the data stream. 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 magnitudes 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. From the energy compaction property of wavelet-transformed image, in highpass subbands, there are more coefficients with small magnitude than those with large magnitudes, as shown in Fig. 7. This implies that more significant bit-planes would contain a lot more 0's than 1's as shown in Fig. 8. Accordingly, the adaptive arithmetic coder would generate more accurate local probability distributions in which the conditional probabilities for "0" symbols are close to 1 for the more significant bit-planes. The context used to determine the conditional probability model of the significant coefficient at $[x, y]$ is related to the significance status of its parent and its eight neighbors. Let $K_p[x, y]$ denote the significance status of the parent, i.e., $K_p[x, y] = 1$ if the parent is significant, otherwise $K_p[x, y] = 0$. Let $K_n[x, y]$ denotes the number significant coefficients in a 3×3 causal neighborhood of the current pixel $[x, y]$. The adaptive context $K[x, y]$ is selected by $K[x, y] = K_n[x, y] + 9K_p[x, y]$, which yields a total of 18 possible models.

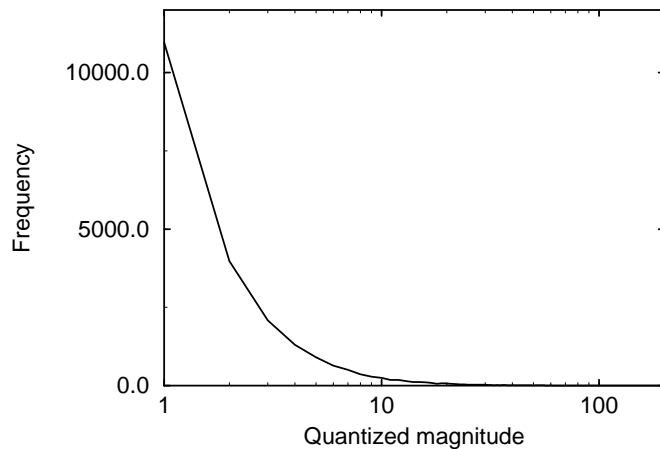


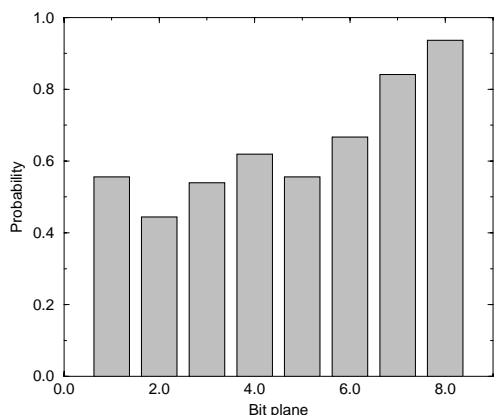
Figure 7: Distribution of significant coefficients of “Lena” image when quantized for final rate 0.5 bpp.

Another method of encoding the coefficients magnitudes is to encode different magnitude value as different source symbols in entropy coding. Experiments have been conducted to compare this method to the bit-plane coding scheme and the results are shown in Table 2. In all cases, bit-plane coding saves 271 to 1264 bytes over direct magnitude encoding. The reason is that for direct magnitude coding, the large number of source symbols demands much more updating steps for the coder to acquire a good estimate of the source distribution compared to bit-plane coding where only two-symbol alphabet is used. As a result, bit-plane magnitude coding exploits the local source distribution more effectively than direct magnitude coding, resulting in higher compression ratio.

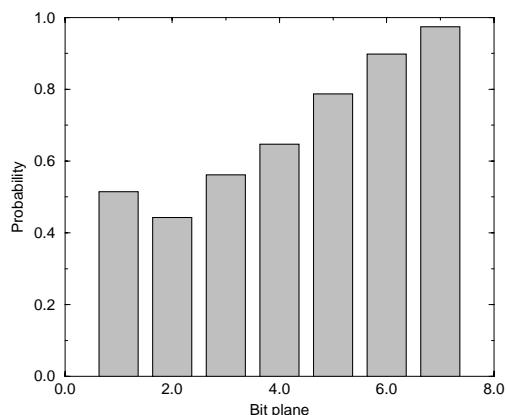
Algorithm/PSNR [dB]	34.33	35.14	36.42	37.38	40.44
Bit-plane coding	8193	9828	13107	16383	32769
Direct magnitude coding	8464	10151	13575	16992	34033

Table 2: Compressed file size [bytes] comparison between bit-plane coding and direct magnitude coding.

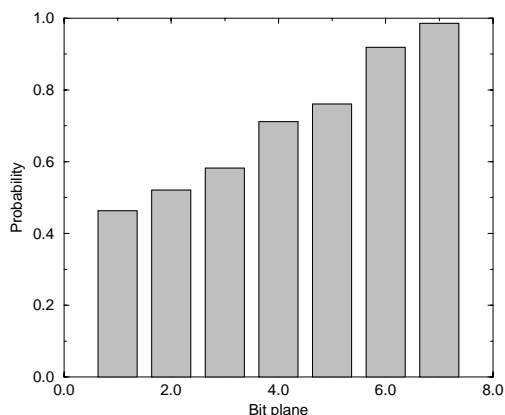
The idea of bit-plane encoding is also used in both EZW and SPIHT, but in a different manner. In EZW, the idea is realized through progressive transmission of magnitudes,



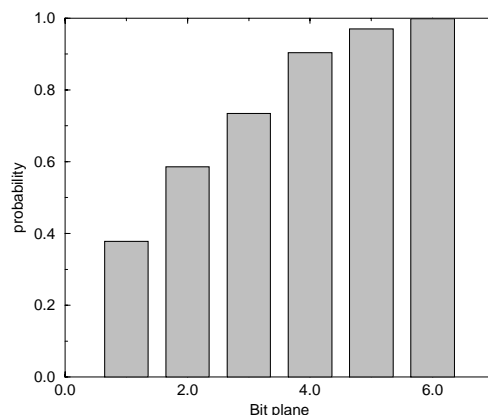
(a) 3rd subband



(b) 5th subband



(c) 7th subband



(d) 12th subband

Figure 8: Zero symbol probabilities in four subbands for “Lena.”

with the “0” bits before the first “1” bit being encoded as either “zerotree” or “isolated zero.” Similar to EZW, in SPIHT, initial “0” bits are represented as part of the insignificant set until the occurrence of the first “1” symbol. Then the magnitudes are coded by progressive transmission.

3.4 Description of SLCCA

In the following, we present the encoding algorithm of SLCCA.

Step 1. Form a subband pyramid and quantize all wavelet coefficients with a uniform scalar quantizer.

Step 2. Perform connected component analysis of significant coefficients within each subband and remove extremely small connected components.

Step 3. Form a scan list containing all the coefficient positions in the subband pyramid as follows. Starting from the coarsest subband, scan subbands according to the order LL, LH, HL, HH (Fig. 4). Within each subband, scan the coefficients from left to right, top to bottom. Go to the next finer scale if all coefficients in the current scale have been scanned.

Step 4. Let c be the first coefficient in the scan list.

Step 5. If c is found significant and has not been encoded, go to **Step 6**; or else, let c be the next coefficient in the scan list and repeat **Step 5**.

Step 6. Encode the position $[x, y]$ of c .

Step 7. Encode the sign (POS or NEG) of c .

Step 8. If c is the parent of a child cluster that has not been linked to any other coefficient, then

8.1 encode a special symbol (LINK);

8.2 move the child position in a first-in-first-out (FIFO) queue to store the information that the child cluster has been linked.

Step 9. For every $[\Delta x, \Delta y]$ in a predefined neighborhood, do

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

then $x = x + \Delta x, y = y + \Delta y$, let c be the coefficient at $[x, y]$ and

go to **Step 7**.

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

then encode a ZERO symbol.

Step 10. If the FIFO queue is not empty, take the next child position out of the queue, and go to **Step 7**, otherwise let c be the next coefficient in the scan list and go to **Step 5**.

Step 11. Encode the magnitude of significant coefficients in bit-plane order using the adaptive arithmetic coder.

The decoding algorithm is straightforward and can be obtained by reversing the encoding process.

4 Performance Evaluation

4.1 Natural Image Compression

The SLCCA is evaluated on eight natural 512×512 grayscale images shown in Fig. 9. 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 10/18 filters (Table 3) obtained from ftp.cs.dartmouth.edu. No optimal bit allocation is carried out in SLCCA. Instead, all wavelet coefficients are quantized with the same uniform scalar quantizer. As usual, the distortion is measured by PSNR defined as

$$\text{PSNR [dB]} = 20 \log_{10} \frac{255}{\text{RMSE}}, \quad (1)$$

where RMSE is the root mean-squared error between the original and reconstructed images. All the reported bit rates are computed from the actual file sizes.

Table 4 shows the comparison among four wavelet coders on “Lena” image at different bit rates. SLCCA consistently outperforms EZW, MRWD, and SPIHT. Compared to EZW, SLCCA gains 1.08 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 on average.

Table 5 compares the performance of SLCCA, EZW, and SPIHT on the “Barbara” image. There exist many versions of “Barbara” image. The one used in this experiment was obtained from ftp.cs.dartmouth.edu which is the same as that used by EZW and



Figure 9: 512×512 test images. From left to right, top to bottom: “Lena,” “Barbara,” “Baboon,” “Couple,” “Man,” “Boat,” “Tank,” and “Goldhill.”

Tap	Value
1	2.885256e-02
2	8.244478e-05
3	-1.575264e-01
4	7.679048e-02
5	7.589077e-01
6	7.589077e-01
7	7.679048e-02
8	-1.575264e-01
9	8.244478e-05
10	2.885256e-02

(a)

Tap	Value
1	9.544158e-04
2	-2.727196e-06
3	-9.452462e-03
4	-2.528037e-03
5	3.083373e-02
6	-1.376513e-02
7	-8.566118e-02
8	1.633685e-01
9	6.233596e-01
10	6.233596e-01
11	1.633685e-01
12	-8.566118e-02
13	-1.376513e-02
14	3.083373e-02
15	-2.528037e-03
16	-9.452462e-03
17	-2.727196e-06
18	9.544158e-04

(b)

Table 3: Biorthogonal filter bank used in SLCCA. (a) Analysis and (b) synthesis filters.

SPIHT. On average, SLCCA is superior to EZW by 1.67 dB, and to SPIHT by 0.79 dB. The original “Barbara” image, and the reconstructed images at 0.25 bpp, 0.5 bpp, and 1.0 bpp are shown in Figs. 10a-d, respectively.

The comparison between SLCCA and SPIHT on the rest of test images is shown in Table 6. SLCCA consistently outperforms SPIHT. It appears that SLCCA performs significantly better than SPIHT for images which are rich in texture; see, for instance, the results of “Barbara,” “Baboon,” “Boat,” and “Tank.” For images which are relatively smooth, the performance between SLCCA and SPIHT gets closer, as indicated by the results of “Goldhill,” “Couple,” and “Man.”

Algorithm/Rate [bpp]	0.125	0.25	0.30	0.40	0.50	1.00
EZW	30.23	33.17	-	-	36.28	39.55
MRWD	-	-	34.07	-	36.60	40.17
SPIHT	31.09	34.11	34.95	36.24	37.21	40.41
SLCCA	31.38	34.33	35.14	36.42	37.38	40.44

Table 4: Performance comparison (PSNR [dB]) on 512×512 “Lena” image.

Algorithm/Rate [bpp]	0.125	0.25	0.30	0.40	0.50	1.00
EZW	24.03	26.77	-	-	30.53	35.14
SPIHT	24.86	27.58	28.56	30.10	31.39	36.41
SLCCA	25.45	28.43	29.39	30.93	32.28	37.15

Table 5: Performance comparison (PSNR [dB]) on 512×512 “Barbara” image.

Rate [bpp]	Algorithm	Image					
		“Baboon”	“Couple”	“Man”	“Boat”	“Tank”	“Goldhill”
0.25	SPIHT	23.27	29.25	30.01	30.97	29.36	30.56
	SLCCA	23.44	29.38	30.07	31.09	29.44	30.64
0.3	SPIHT	23.76	30.07	30.74	31.77	29.77	31.15
	SLCCA	23.99	30.10	30.81	31.94	29.89	31.24
0.4	SPIHT	24.66	31.29	31.94	33.16	30.52	32.18
	SLCCA	24.96	31.43	32.04	33.44	30.65	32.32
0.5	SPIHT	25.64	32.45	33.08	34.45	31.16	33.13
	SLCCA	25.83	32.55	33.12	34.68	31.28	33.24
1.0	SPIHT	29.17	36.58	37.34	39.12	33.78	36.55
	SLCCA	29.33	36.58	37.33	39.28	33.99	36.66

Table 6: Performance comparison (PSNR [dB]) of SPIHT and SLCCA on different 512×512 natural images.



(a)



(b)



(c)



(d)

Figure 10: Coding results for 512×512 “Barbara” image. (a) Original, reconstructed images (b) at 1.0 bpp, PSNR=37.15 dB, (c) at 0.5 bpp, PSNR=32.28 dB, and (d) at 0.25 bpp, PSNR=28.43 dB.

4.2 Texture Image Compression

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. 11. The results at 0.4 bpp are summarized in Table 7 indicating that SLCCA constantly outperforms SPIHT by 0.32 dB to 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 7: Performance comparison (PSNR [dB]) of SPIHT and SLCCA on 256×256 texture images at 0.4 bpp.

4.3 Comparison with the Latest Wavelet Coding Algorithms

Recently, we learned about two new wavelet image coding algorithms, i.e., Xiong *et al.*'s wavelet-based space-frequency quantization [23] (SFQ), and a latest version of Servetto *et al.*'s MRWD [24].

In SFQ, the zerotree structure is optimized for a given target bit rate using the Lagrange multiplier method in the operational rate-distortion sense. The optimization

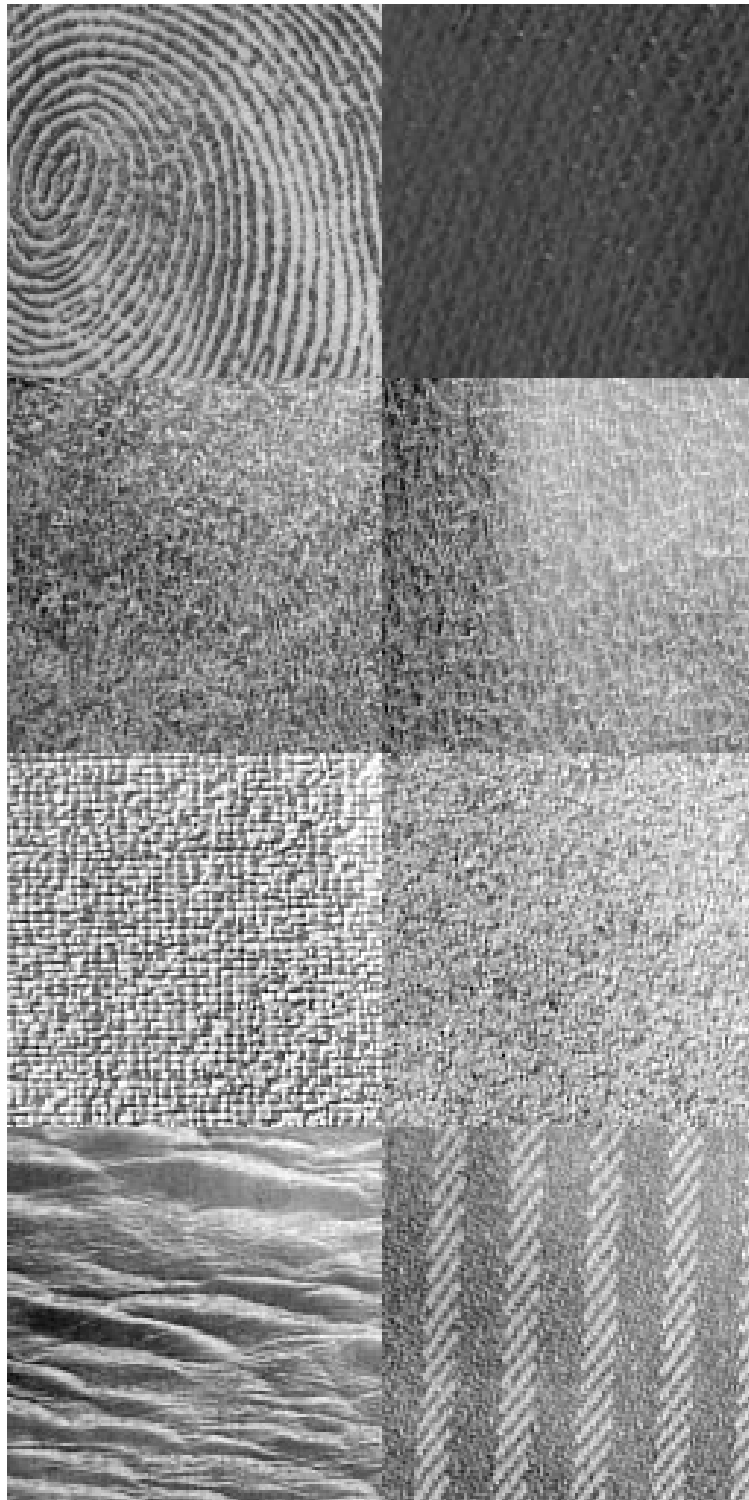


Figure 11: 256×256 texture images. From left to right, top to bottom: “fingerprint,” “sweater,” “grass,” “pig skin,” “raffia,” “sand,” “water” and “wool.”

procedure yields remarkable performance, but at the price of much higher computational complexity.

In the latest version of MRWD [24], all wavelet coefficients are coded regardless of their significance. Raster scan is the basic scan order to encode the coefficients in a subband. When a significant coefficient (the seed of a cluster) is encountered, a special symbol is encoded followed by the entire cluster before continuing the raster scan. The context of the adaptive arithmetic model is mainly based on the significance of the parent coefficient. As a single uniform quantizer is used for bit rate control, the complexity of both the encoding and decoding process is comparable to SLCCA.

The performance comparison of the latest MRWD, SFQ and SLCCA is given in Table 8. In the case of the “Lena” image, SFQ and SLCCA yield comparable performance. For the other two test images, SLCCA is superior on “Barbara,” while SFQ is superior on “Goldhill.” The differences in PSNR for both images are about 0.1 dB. Finally, both SLCCA and SFQ outperform MRWD.

Image	Algorithm/Rate [bpp]	0.25	0.30	0.40	0.50	1.00
“Lena”	MRWD [24]	34.12	34.93	36.20	37.18	40.33
	SFQ	<u>34.33</u>	35.07	<u>36.43</u>	37.36	<u>40.52</u>
	SLCCA	<u>34.33</u>	<u>35.14</u>	36.42	<u>37.38</u>	40.44
“Barbara”	MRWD [24]	27.86	28.71	30.22	31.44	36.24
	SFQ	28.29	29.21	30.77	32.15	37.03
	SLCCA	<u>28.43</u>	<u>29.39</u>	<u>30.93</u>	<u>32.28</u>	<u>37.14</u>
“Goldhill”	MRWD [24]	30.53	31.14	32.19	33.15	36.56
	SFQ	<u>30.71</u>	<u>31.34</u>	<u>32.45</u>	<u>33.37</u>	<u>36.70</u>
	SLCCA	30.64	31.24	32.32	33.25	36.66

Table 8: Performance comparison of the latest MRWD, SFQ and SLCCA on “Lena,” “Barbara,” and “Goldhill.”

Image	“Foreman”		“Hall Monitor”	
Rate [bits]/ Algorithm	14122	27132	14164	28280
H.263	31.62	36.82	31.42	37.57
SLCCA	32.69	37.57	32.04	38.34

Table 9: Performance comparisons of I-frame coding results from H.263 and SLCCA.

4.4 SLCCA for Intraframe Coding in Video Compression

In most video coding algorithms, the frames in a video sequence are divided into two categories: intraframe and interframe. Intraframes are coded using still image compression algorithms, while interframes are predicted from intraframes or other interframes. In the existing international standards for video compression, such as H.261, H.263, MPEG-1, and MPEG-2, intraframes are encoded using block-based DCT coding algorithms similar to JPEG. Tables 9 compare the results of encoding the intraframe by both H.263 and SLCCA. Figs. 12 and 13 show the reconstructed frames from both algorithms for the “Foreman” and “Hall Monitor” sequences, respectively. It is clear that for intraframe compression, SLCCA outperforms H.263 in both objective and subjective measure and thus could be a very good candidate for intraframe compression.

5 Conclusions

A new image coding algorithm termed significance-linked connected component analysis is proposed. 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 show that SLCCA is among the best image coding algorithms reported in the literature.



(a)



(b)



(c)

Figure 12: Coding results. (a) Original 1st frame of “Foreman” sequence. Reconstructed images from (b) H.263 and (c) SLCCA at 14 kbits/frame.



(a)



(b)



(c)

Figure 13: Coding results. (a) Original 1st frame of “Hall Monitor” sequence. Reconstructed images from (b) H.263 and (c) SLCCA at 14 kbits/frame.

Acknowledgment

The authors would like to thank the reviewer for his comments and suggestions. The surface plots of Fig. 2 were generated by Interactive Image SpreadSheet (IISS) [25] developed at NASA Goddard Space Flight Center.

References

- [1] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
- [2] O. Rioul and M. Vetterli, “Wavelets and signal processing,” *IEEE Signal Processing Magazine*, pp. 14–38, Oct. 1991.
- [3] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [4] J.W. Woods, Ed., *Subband Image Coding*, Kluwer Academic Publishers, Norwell, MA, 1991.
- [5] 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.
- [6] S.G. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, July 1989.
- [7] J.W. Woods and S.D. O’Neil, “Subband coding of images,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 5, pp. 1278–1288, Oct. 1988.
- [8] 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.
- [9] G.K. Wallace, “The JPEG still picture compression standard,” *Communications of ACM*, vol. 34, no. 4, pp. 30–44, Apr. 1991.
- [10] 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.

- [11] 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.
- [12] 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.
- [13] J.B. Allen and L.R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proceedings of IEEE*, vol. 65, pp. 1558–1564, 1977.
- [14] A.S. Lewis and G. Knowles, "A 64 Kb/s video codec using the 2-D wavelet transform," in *Proceedings of Data Compression Conference*, Snowbird, Utah, 1991.
- [15] B.-B. Chai, J. Vass, and X. Zhuang, "Significance-linked connected component analysis for wavelet image coding," *IEEE Transactions on Image Processing*, submitted, 1997.
- [16] X. Li and X. Zhuang, "The decay and correlation properties in wavelet transform," Tech. Rep., University of Missouri-Columbia, Mar. 1997.
- [17] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and effective algorithms," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 176–201, Apr. 1993.
- [18] R.M. Haralick, S.R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 532–550, July 1987.
- [19] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, 1992.
- [20] J. Luo, C.W. Chen, K.J. Parker, and T.S. Huang, "A scene adaptive and signal adaptive quantization for subband image and video compression using wavelets," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 343–357, Apr. 1997.
- [21] 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.
- [22] 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.

- [23] Z. Xiong, K. Ramchandran, and M.T. Orchard, “Space-frequency quantization for wavelet image coding,” *IEEE Transactions on Image Processing*, vol. 6, no. 5, pp. 677–693, May 1997.
- [24] S. Servetto, K. Ramchandran, and M.T. Orchard, “Image coding based on morphological representation of wavelet data,” *IEEE Transactions on Image Processing*, *submitted*, 1996.
- [25] A.F. Hasler, K. Palaniappan, M. Manyin, and J. Dodge, “A high performance interactive image spreadsheet (IIS),” *Computers in Physics*, vol. 8, no. 3, pp. 325–342, 1994.