Significance-Linked Connected Component Analysis for Wavelet Image Coding

Bing-Bing Chai *Jozsef Vass* *Xinhua Zhuang*

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

Abstract

Recent success in wavelet image coding is mainly attributed to recognition of the importance of data organization and representation. There have been several very competitive wavelet coders 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 extends MRWD by exploiting both within-subband clustering of significant coefficients and cross-subband dependency in significant fields. Extensive computer experiments on both natural and texture images show convincingly that the proposed SLCCA outperforms EZW, MRWD, and SPIHT. For example, for the "Barbara" image, at 0.25 bpp SLCCA outperforms EZW, MRWD and SPIHT by 1.41 dB, 0.32 dB and 0.60 dB in PSNR, respectively. It is also observed that SLCCA works extremely well for images with a large portion of texture. For eight typical 256×256 grayscale texture images compressed at 0.40 bpp, SLCCA outperforms SPIHT by 0.16 dB–0.63 dB in PSNR. This outstanding performance is achieved without using any optimal bit allocation procedure. Thus both the encoding and decoding procedures are fast.

Submitted: IEEE Trans. on Image Processing
EDICS category: IP 1.1 (Coding)
Correspondence author: Dr. Xinhua Zhuang
Address: 313 Engineering Building West
Department of Computer Engineering & Computer Science
University of Missouri-Columbia
Columbia, MO 65211
Phone: (573) 882-2382
Fax: (573) 882-8318
E-mail: zhuang@cecs.missouri.edu

1 Introduction

Since its introduction for speech coding in the 70's, subband coding [1] has become a very active research area for image and video compression. Wavelet theory [2, 3, 4] provides a fundamental insight into the structure of subband filters that leads to a more productive approach to designing the filters [2, 5]. This is evidenced by the discovery of symmetric biorthogonal wavelet bases with compact support [2, 6] which are instantly converted into more desirable linear phase filters while maintaining the necessary perfect reconstruction. Thus subband and wavelet are often used interchangeably in the literature.

Most of the subband image coders published recently are based on pyramidal (or dyadic) wavelet decomposition as shown in Fig. 1. Conventional wavelet or subband image coders [6, 7] 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 [8], their objective performance measured by PSNR increases only moderately.

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 statistical properties in a wavelet pyramid one way or the other. There are three representatives of such top-ranked wavelet image coders, namely, Shapiro's embedded zerotree wavelet coder (EZW) [9], Servetto *et al.*'s morphological representation of wavelet data (MRWD) [10, 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 propose a novel strategy for data organization and representation for

wavelet image coding termed significance-linked connected component analysis (SLCCA). 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 decomposition, connected component analysis of significant fields within subbands, and significance-link registration across subbands, as well as bit plane encoding of magnitudes of significant coefficients by adaptive arithmetic coding.

The rest of the paper is organized as follows. In Section 2, we discuss the statistical properties of wavelet-transformed images. In Section 3, we analyze and compare the data organization and representation strategies by EZW, MRWD, and SPIHT. Our wavelet image coding algorithm, SLCCA, is presented in Section 4. Section 5 presents the performance evaluation. In Section 5.1, the performance of SLCCA is evaluated against other wavelet coders. Performance comparison among SLCCA and recently published most state-of-the-art codecs is given in Section 5.2. The last section concludes the paper.

2 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.

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

Each wavelet coefficient contains only features from a local segment of an input image, i.e., it is spatially localized. 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.

A natural image is typically composed of a large portion of homogeneous and textured regions, together with 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 mostly 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 using only one fourth the size of the original lowpass subband. Repeating this decomposition process on an image will effectively compacted the energy into few wavelet 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. 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 indicate the occurrence of edges or textures with 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 Figs. 1a and 2a.

Relative to a given wavelet coefficient, all coefficients at finer scales of similar orientation which correspond to the same spatial location are called its *descendents*; 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. 1b). These concepts were introduced by Lewis and Knowles in [13]. Although the linear correlation between the values of parent and child wavelet coefficients has been empirically found to be extremely small, there is additional

dependency between the magnitudes of parent and children. Experiments showed that the correlation 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 [9].

Although it appears to be difficult to characterize and make a full use of this cross-subband magnitude similarity, a reasonable conjecture based on experience with real-world images is that the magnitude of a child is 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 decays from the parent to its children. More precisely, if we measure the coefficient magnitude by its variance, we can prove that a parent has larger variance than its children [14]. This provides a strong theoretical support to EZW, SPIHT, and SLCCA.

3 Overview of Data Organization and Representation Strategies

There exist two different approaches to an efficient organization and representation of wavelet coefficients in literature [9, 12, 11]. While EZW and SPIHT use a regular tree structure or set-partitioned tree structure to approximate insignificant fields across subbands, MRWD finds irregular clusters of significant fields within subbands.

It is widely 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 zero regions across subbands by the union of highly constrained tree-structured zero regions called *zerotrees*. Meanwhile, it defines the significant fields everywhere 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; the last part

comprising 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 all the insignificant non-child descendents as a set and employs a single symbol to represent and encode it. This fine set partitioning strategy leads to an impressive increase in PSNR by 0.86–0.94 dB over EZW on "Lena" image (see Table 1), indicating that SPIHT exploits cross-subband dependency more efficiently than EZW.

Different from EZW and SPIHT, MRWD [10, 11] 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 [15] which utilizes a structuring element to control the shape and size of clusters as well as the formation of boundaries. Details on cluster formation will be presented in the following section. With MRWD, the boundary zeros of each cluster still need to be represented 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 0.78–0.95 dB over EZW on "Lena" image, as shown in Table 1. Nevertheless, in the early version of MRWD [10], a seed for each cluster, i.e., a pixel from which a cluster is originated, needs to be specified and its positioning information is encoded as overhead. Since a large number of clusters are involved, the overall overhead may take up a significant portion of the bit budget. Our new coding algorithm is developed based on our knowledge of this early version of MRWD. In its latest version [11], MRWD keeps its main feature, i.e., the within-subband clustering. However, the coding method has been changed. All wavelet coefficients are coded regardless of their significance. The seed of each cluster is specified by transmitting a special symbol. In addition, a context based adaptive arithmetic model is employed for entropy coding, where the context is based on the significance of the parent coefficient. The use of this adaptive arithmetic model is an *implicit* exploitation of the cross subband similarity which is expected to improve the compression. The experimental results shows that the latest MRWD indeed outperforms the early version (Table 1).

4 Significance-Linked Connected Component Analysis

In this section, the key features of our wavelet coder SLCCA is first described. Then a complete algorithm is presented.

4.1 Formation of Connected Components within Subbands

First, we will review some morphological operations relevant to our algorithm. More detailed discussion of mathematical morphology can be found in [15, 16, 17]. A binary image can be considered as a subset of $E \times E$, where E denotes the set of numbers used to index a row or column position on a binary image. Pixels are in this subset if and only if they have the binary value "1" on the image. The dilation of set $A \subseteq E^2$ with set $B \subseteq E^2$ is defined by $A \oplus B = \cup_{b \in B} A_b$, where A_b denotes the translation of A by b, $\cup$ the pixel-wise union. For a structuring element that contains the origin (such as those used in SLCCA and shown in Fig. 3), the dilation operation produces an enlarged set containing the original set A and some neighboring pixels. Let $A \subseteq E^2$ denote the set to be reconstructed and the marker $M \subseteq A$ be an arbitrary subset of A. Then the *conditioned dilation* operation is defined as follows [15]:

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

where $\cap$ is the pixel-wise intersection. If we let

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

then a cluster is formed when $D^n(M, A) = D^{n-1}(M, A)$.

Since a rather large portion of wavelet coefficients are usually insignificant and significant coefficients within subbands tend to be more clustered, organizing and representing each subband as irregular shaped clusters of significant coefficients provide 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 [10]. In the following, we discuss the issue in regard to the selection of structuring elements.

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 $M \subset A$ contains 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 a 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. 3. The ones in Figs. 3a and 3b generate 4- and 8-connectivity, respectively. The structuring elements in Figs. 3c and 3d represent a diamond of size 13 and a 5×5 square, respectively. These latter two may not preserve geometric connectivity but perform better than the former in terms of coding efficiency.

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, which will be presented in the next subsection, largely reduces the positioning cost, relatively smaller structuring elements can be selected for connected component analysis.

The operation of progressive cluster detection by conditioned dilation operation is illustrated in Fig. 4, where image size of 5×5 is assumed and 4-connected structuring element shown in Fig. 3a is used. The significant map A is shown in Fig. 4a. In the example, pixel at (3,2) is chosen as seed, i.e., $M = \{(3,2)\}$, and the remaining 7 steps of the recursive cluster detection are shown in Figs. 4c–4i.

As extremely small clusters usually do not produce discernible visual effects, and these clusters render a higher insignificant-to-significant coefficient ratio than large clusters, they are eliminated to avoid 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. There are 22748 significant wavelet coefficients after quantization, forming 1654 clusters using the structuring element shown in Fig. 3c. 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. 2c with the marker image shown in Fig. 2b. It is clear that only a small fraction of zero coefficients are encoded.

4.2 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 second generation descendents in SPIHT are more of a reality when a large threshold is used. As was stated in [9] and [18], when the threshold decreases (for embedding) to a certain point, the tree structure or set-partitioned-tree structure is 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.

Now, we define *significance-link* 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. 5). 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. An experiment has been conducted to test the effectiveness of significance-link. Two versions of the algorithm are tested under the condition that all parameters are set the same except for that one version uses significance-link while the other does not. It has been shown that the saving from using significance-link over without 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.

4.3 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 [19] is employed. In contrast to a fixed model arithmetic coder, which works well for a stationary Markov source, the adaptive arithmetic coder updates the corresponding 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 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 [20], 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, there are more coefficients with small magnitudes than those with large magnitudes, implying that the 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 close to one for the more significant bit-planes.

The context used to determine the conditional probability model of 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 pixel 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.

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."

4.4 Description of SLCCA Algorithm

In the following, we summarize the previous three subsections with the encoding algorithm of SLCCA. Four symbols are used to encode the shape of clusters: POS, NEG, ZERO, and LINK. POS or NEG represents the sign of a significant coefficient. ZERO represents an insignificant

coefficient that delineates the boundary of a cluster. LINK marks the presence of a significance-link. The magnitudes of significant coefficients are encoded in bit-plane order with two symbols: "0" and "1." Three lists of coefficients are maintained in the algorithm: list of scan order (LSO), list of child clusters (LCC), and list of significant coefficients (LSC). All these lists are first-in-first-out queues. Each entry in the lists is identified by a coordinate $[x, y]$. $c[x, y]$ denotes the coefficient at position $[x, y]$.

BEGIN **SLCCA-encode()**

Step 1. Form a subband pyramid and quantize all wavelet coefficients with a uniform scalar quantizer. The quantization step size is selected such that the target bit rate is satisfied.

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

Step 3. Form LSO containing all coefficient positions in the subband pyramid as follows. Starting from the coarsest subband, scan subbands according to the order LL, LH, HL, HH (Fig. 1b). Within each subband, scan the coefficients from left to right, top to bottom. Go to the next finer scale after all coefficients in the current scale have been scanned.

Step 4. Encoding clusters.

4.1. Start of a new cluster. For every entry $[x, y]$ in LSO, if $c[x, y]$ is significant and has not yet been encoded,

4.1.1. Encode the position $[x, y]$, i.e., $c[x, y]$ is the seed of a cluster.

4.1.2. Call **Encode-significant-coeff**(x, y).

4.2. Encode child clusters. For each entry $[x, y]$ in LCC,

4.2.1. Remove $[x, y]$ from LCC.

4.2.2. For $\Delta x = 0, 1$, and $\Delta y = 0, 1$,

4.2.2.1. If $c[x + \Delta x, y + \Delta y]$ is significant and has not been encoded, goto **Step 4.2.3**.

4.2.2.2. If $c[x+\Delta x, y+\Delta y]$ is insignificant and has not been encoded, encode a ZERO symbol.

4.2.3. Call **Encode-significant-coeff**(x, y).

Step 5. Encode the magnitude of significant coefficients, i.e., all the entries from the LSC list, in bit-plane order using the adaptive arithmetic coder.

END **SLCCA-encode()**

BEGIN **Encode-significant-coeff**(x, y)

Step 1. Encode the sign (POS or NEG) of $c[x, y]$, put the position $[x, y]$ to the end of LSC.

Step 2. If $c[x, y]$ is the parent of a child cluster that has not been linked to any other coefficient, then

2.1. Encode a special symbol (LINK);

2.2. Move the child position $[2x, 2y]$, which represents all four children of $c[x, y]$, to the end of LCC. This indicates that the child cluster has been linked.

Step 3. Expanding a cluster. For every $[\Delta x, \Delta y]$ in a predefined neighborhood,

3.1. if $c[x+\Delta x, y+\Delta y]$ is significant and has not been encoded, then call **Encode-significant-coeff**$(x+\Delta x, y+\Delta y)$.

3.2. if $c[x+\Delta x, y+\Delta y]$ is insignificant, then encode a ZERO symbol.

END **Encode-significant-coeff**(x, y)

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

5 Performance Evaluation

5.1 Comparison of Algorithms Using Different Data Organization Strategies

SLCCA is evaluated on eight natural 512×512 grayscale images, i.e., "Lena," "Barbara," "Baboon," "Couple," "Man," "Boat," "Tank," and "Goldhill." The performance is compared with

three wavelet coders EZW, MRWD, and SPIHT. In SLCCA, each original image is decomposed into a six-scale subband pyramid using the 9/7 biorthogonal filters [6]. There is no optimal bit allocation carried out in SLCCA. Instead, all wavelet coefficients are quantized with the same uniform scalar quantizer. As usual, the distortion is measured by peak signal-to-noise ratio (PSNR) defined as

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

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.

MRWD, SPIHT, and SLCCA all use the same 9/7 biorthogonal filters and six-scale dyadic wavelet decomposition. EZW also uses six-scale dyadic wavelet decomposition. However, a somewhat older filter proposed in [21] is used to obtain wavelet transform.

Table 1 shows the PSNR comparison on "Lena" image at different bit rates. SLCCA consistently outperforms EZW, both versions of MRWD, and SPIHT as well. Compared to EZW, SLCCA gains 1.03 dB in PSNR on average. When compared to the latest version MRWD [11], SLCCA is superior by 0.16 dB on average. Compared to SPIHT, SLCCA gains 0.13 dB on average.

Table 2 compares the performance of SLCCA, EZW, MRWD and SPIHT on "Barbara" image. On average, SLCCA is superior to EZW by 1.42 dB, and to SPIHT by 0.47 dB. SLCCA also outperforms MRWD [11] by 0.33 dB on average. The original "Barbara" image, and the reconstructed images at 0.125 bpp, 0.25 bpp, and 0.5 bpp are shown in Figs. 6a-d, respectively.

The comparison between SLCCA, MRWD and SPIHT on the rest of test images is shown in Table 3. SLCCA consistently outperforms both SPIHT and MRWD. 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." A similar observation is also true for MRWD [11], i.e., for the texture-rich images, MRWD outperforms SPIHT in general, indicating that the use of clustering is superior to zerotree structure for texture images.

To further verify the above observation, we compare the performance of SLCCA and SPIHT on eight typical 256×256 grayscale texture images, i.e., "fingerprint," "sweater," "grass," "pig skin," "raffia," "sand," "water," and "wool," as shown in Fig. 7. The results at 0.4 bpp are summarized in Table 4, indicating that SLCCA constantly outperforms SPIHT by 0.16 dB to 0.63 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.

Finally, we apply SLCCA to fingerprint image compression, which represents a very important issue demanding the best solution. As known, the digitized fingerprints of a person may require 10 Mbyte of storage without any compression. With such a huge amount of data, the real time transmission of uncompressed fingerprints becomes impossible. The FBI has developed a fingerprint image compression algorithm called wavelet scalar quantization (WSQ) [22]. Table 5 lists the coding results of the 768×768 "Fingerprint" image from WSQ, SPIHT, and SLCCA. Again, SLCCA outperforms SPIHT by an average of 0.26 dB. At 0.444 bpp or 18:1 compression, SLCCA yields a PSNR of 35.65 dB as opposed to WSQ's 34.43 dB, corresponding to a 1.22 dB improvement. The original and reconstructed images from SLCCA at 0.444 bpp are shown in Fig. 8.

The coding results along with the images are also available at the homepage of the Multimedia and Visualization Laboratory at http://meru.cecs.missouri.edu.

5.2 Comparison with the State-of-the-Art

To the best of our knowledge, there have been three other very competitive wavelet coders proposed recently, namely, Xiong *et al.*'s space-frequency quantization (SFQ) [23], LoPresto *et al.*'s estimation quantization (EQ) [24], and Joshi *et al.*'s optimal classification (OC) [25].

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

remarkable performance at the price of high computational complexity.

In EQ, the wavelet coefficients is modeled as a generalized Gaussian distribution with zero mean and unknown variance. For each subband, the variance of wavelet coefficients is estimated by Maximum Likelihood estimator. Then, each wavelet coefficient is quantized with an off-line designed optimal quantizer scaled to match the estimated variance of wavelet field. The distinct feature of EQ is the backward adaptive magnitude estimation and quantization of a wavelet coefficient based on its quantized neighboring coefficients.

In OC, the variance of coefficients in each subband is estimated by an iterative algorithm, after which small blocks of coefficients are classified into a given number of classes based on their variance. Then, each class is modeled by a Generalized Gaussian density and optimal bit allocation is carried out among classes from all the subbands.

The performance evaluation is given in Table 6. The performance comparison between the above mentioned coding algorithms and SLCCA is fairly difficult due to the different type of filter and wavelet decomposition used by different algorithm.

SLCCA and SFQ both use the same wavelet filter and dyadic wavelet decomposition. The performance of SLCCA and SFQ for the "Lena" and "Goldhill" images is comparable, i.e., SFQ slightly outperforms SLCCA by 0.03 dB and 0.08 dB on average, for the "Lena" and "Goldhill" images, respectively. For the "Barbara" image SFQ exceeds SLCCA by 0.22 dB on average. EQ is superior to SLCCA by 0.35 dB and 0.21 dB on average for the "Lena" and "Goldhill" images, respectively. Nevertheless, EQ uses 4-scale dyadic wavelet transform with 10/18 normalized biorthogonal filter set [26] which gives a slightly superior performance when compared to 9/7 filter used in SLCCA. Finally, OC uses a 22-band decomposition. For the "Lena" image at 0.25 bpp, SLCCA is slightly superior to OC by 0.03 dB, and at 0.5 bpp and 1.0 bpp OC exceeds SLCCA by 0.28 and 0.83 dB, respectively. Unlike SLCCA, the exact bit rate control of OC is not solved. Thus at 0.25 bpp and 1.0 bpp the actual rate by OC exceeds the target rate.

6 Conclusions

A new image coding algorithm termed significance-linked connected component analysis is proposed 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 SLCCA is among the state-of-the-art image coding algorithms reported in the literature. As no optimization is involved, both the encoding and decoding procedures are fast.

Acknowledgment

The authors would like to thank the reviewers for their valuable comments and suggestions.

References

[1] J.W. Woods, Ed., *Subband Image Coding*, Kluwer Academic Publishers, Norwell, MA, 1991.

[2] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[3] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, pp. 14–38, Oct. 1991.

[4] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[5] 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.

[6] 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.

[7] 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.

[8] G.K. Wallace, "The JPEG still picture compression standard," *Communications of ACM*, vol. 34, no. 4, pp. 30–44, Apr. 1991.

[9] 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.

[10] 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.

[11] S. Servetto, K. Ramchandran, and M.T. Orchard, "Image coding based on morphological representation of wavelet data," *IEEE Transactions on Image Processing, submitted*, 1996.

[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] 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.

[14] X. Li and X. Zhuang, "The decay and correlation properties in wavelet transform," Tech. Rep., University of Missouri-Columbia, Mar. 1997.

[15] 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.

[16] 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.

[17] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, 1992.

[18] 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.

[19] 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.

[20] M. Rabbani and P.W. Jones, *Digital Image Compression Techniques*, SPIE Optical Engineering Press, Belingham, WA, 1991.

[21] E.H. Adelson and E. Simoncelli, "Orthogonal pyramid transforms for image coding," in *Proceedings of SPIE Conference Visual Communications and Image Processing*, 1987, vol. 845, pp. 50–58.

[22] 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.

[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.M. LoProso, K. Ramchandran, and M.T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework," in *Proceedings of Data Compression Conference*, 1997.

[25] R.L. Joshi, H. Jafarkhani, J.H. Kasner, T.R. Fischer, N. Farvardin, M.W. Marcellin, R.H. Bamberger, "Comparison of different methods of classification in subband coding of images," *IEEE Transactions on Image Processing*, vol. 6, no. 11, pp. 1473–1486, Nov. 1997.

[26] B. Usevitch, "Optimal bit allocation for biorthogonal wavelet coding," in *Proceedings of Data Compression Conference*, Snowbird, Utah, 1996, pp. 387–395.

Biography

Bing-Bing Chai received the B.S. degree in physics from Peking University, Beijing, China, in 1990. She received the M.S. degree in medical physics in 1992 and the Ph.D. degree in electrical engineering in 1997 from the University of Missouri-Columbia.

In 1993-1997, she was a teaching and research assistant in the Department of Electrical and Computer Engineering at University of Missouri-Columbia. In October 1997, she joined the Multimedia Technology Lab in the Sarnoff Corporation, Princeton, New Jersey. Her research interest include video and image compression, multimedia signal processing, digital communication and networking.

Jozsef Vass received the Dipl. Eng. in Electrical Engineering from the Technical University of Budapest, and the M.S. degree in Electrical Engineering from the University of Missouri-Columbia, where he is currently a a Ph.D. candidate at the Department of Computer Engineering and Computer Science.

He worked for NASA Goddard Space Flight Center, Greenbelt, MD in the summer of 1996 in the development of robust algorithms for automatic cloud height estimation. He has over 20 refereed technical journal and conference publications. His research interest include speech, image, and video compression for multimedia communications, networking, computer vision, image processing, and pattern recognition.

Xinhua Zhuang is currently Professor of Computer Engineering and Computer Science at the University of Missouri-Columbia. He received the B.S., M.S., and Ph.D. degrees in Mathematics from the Peking University, China, in 1959, 1960, and 1963, respectively. He has been consultant to Siemens, Panasonic, NeoPath Inc., and NASA. He has been affiliated with a number of schools and research institutes including Hannover University, Germany, Zhejiang University, China, the University of Washington, USA, the University of Michigan, USA, the Virginia Polytechnic Institute and State University of Virginia, USA, and the Research Institute of Computers. Professor Zhuang has over 200 publications in the areas of signal processing, speech recognition, image processing, machine vision, pattern recognition, and neural networks. He is a contributor of seven books.

He has received awards from NSF, NASA High Performance Computing and Communications, NASA Innovative Research and NATO Advisory Group of Aerospace Research and Development (AGARD).

Professor Zhuang joined the IEEE in 1992 as Senior Member and was Associate Editor for IEEE Transactions on Image Processing during 1993-1995. Since 1997, he serves as Chairman of Benchmarking & Software Technique Committee for International Association of Pattern Recognition.

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 [10]	-	-	34.07	-	36.60	40.17
MRWD [11]	31.09	34.12	34.93	36.20	37.17	40.33
SPIHT	31.09	34.11	34.95	36.24	37.21	40.41
SLCCA	31.25	34.28	35.08	36.37	37.35	40.47

Table 1: 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
MRWD [11]	25.27	27.86	28.71	30.22	31.44	36.24
SPIHT	24.86	27.58	28.56	30.10	31.39	36.41
SLCCA	25.36	28.18	29.05	30.55	31.89	36.69

Table 2: Performance comparison (PSNR [dB]) on 512×512 "Barbara" image.

Rate [bpp]	Algorithm	Image					
		"Baboon"	"Couple"	"Man"	"Boat"	"Tank"	"Goldhill"
	SPIHT	23.27	29.25	30.01	30.97	29.36	30.56
0.25	MRWD [11]	23.26	29.19	29.95	30.92	29.42	30.53
	SLCCA	23.44	29.37	30.07	31.09	29.44	30.60
	SPIHT	23.76	30.07	30.74	31.77	29.77	31.15
0.3	MRWD [11]	23.83	29.93	30.69	31.77	29.86	31.14
	SLCCA	23.98	30.12	30.83	31.96	29.89	31.27
	SPIHT	24.66	31.29	31.94	33.16	30.52	32.18
0.4	MRWD [11]	24.84	31.25	31.95	33.20	30.61	32.19
	SLCCA	24.96	31.42	32.08	33.39	30.62	32.36
	SPIHT	25.64	32.45	33.08	34.45	31.16	33.13
0.5	MRWD [11]	25.74	32.36	33.06	34.39	31.26	33.15
	SLCCA	25.86	32.59	33.19	34.61	31.27	33.26
	SPIHT	29.17	36.58	37.34	39.12	33.78	36.55
1.0	MRWD [11]	29.35	36.48	37.26	38.96	33.99	36.56
	SLCCA	29.36	36.64	37.45	39.24	34.04	36.66

Table 3: Performance comparison (PSNR [dB]) of SPIHT, MRWD, and SLCCA on different 512×512 natural images.

Image	SLCCA	SPIHT
"fingerprint"	27.37	27.07
"sweater"	41.84	41.48
"grass"	25.33	24.82
"pig skin"	26.69	26.50
"raffia"	20.93	20.30
"sand"	24.10	23.63
"water"	29.35	29.19
"wool"	26.14	25.70

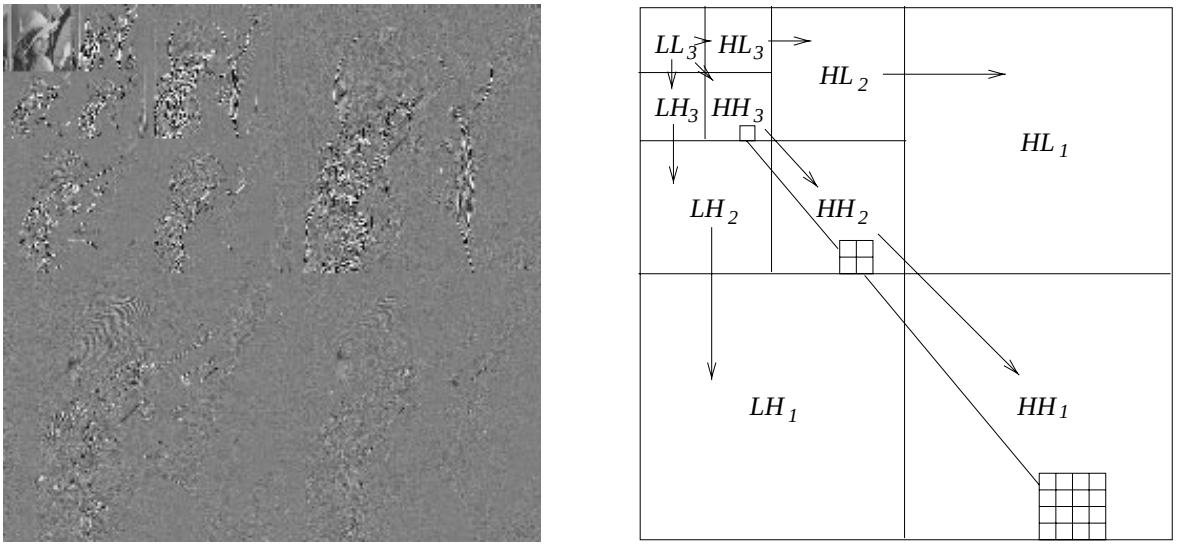
Table 4: Performance comparison (PSNR [dB]) of SPIHT and SLCCA on 256×256 texture images at 0.4 bpp.

Algorithm/Rate [bpp]	0.125	0.25	0.30	0.40	0.444	0.50	1.00
WSQ	-	-	-	-	34.43	-	-
SPIHT	29.46	32.82	33.62	34.98	35.45	36.03	40.00
SLCCA	29.80	33.09	33.89	35.16	35.65	36.25	40.32

Table 5: Performance comparison (PSNR [dB]) on 768×768 "Fingerprint" image.

Image	Algorithm/Rate [bpp]	0.25	0.30	0.40	0.50	1.00
"Lena"	SFQ	34.33	35.07	36.43	37.36	40.52
	OC	34.31	-	-	37.63	41.30
	EQ	34.57	-	-	37.69	40.89
	SLCCA	34.28	35.08	36.37	37.35	40.47
"Barbara"	SFQ	28.29	29.21	30.77	32.15	37.03
	SLCCA	28.18	29.05	30.55	31.89	36.69
"Goldhill"	SFQ	30.71	31.34	32.45	33.37	36.70
	EQ	30.76	-	-	33.42	36.96
	SLCCA	30.60	31.27	32.36	33.26	36.66

Table 6: Performance comparison of the SFQ, EQ, OC, and SLCCA on "Lena," "Barbara," and "Goldhill" images.



(a) (b)

Figure 1: Wavelet pyramid. (a) Three-scale wavelet decomposition for the "Lena" image. (b) Illustration of parent-child relationship between subbands at different scales.

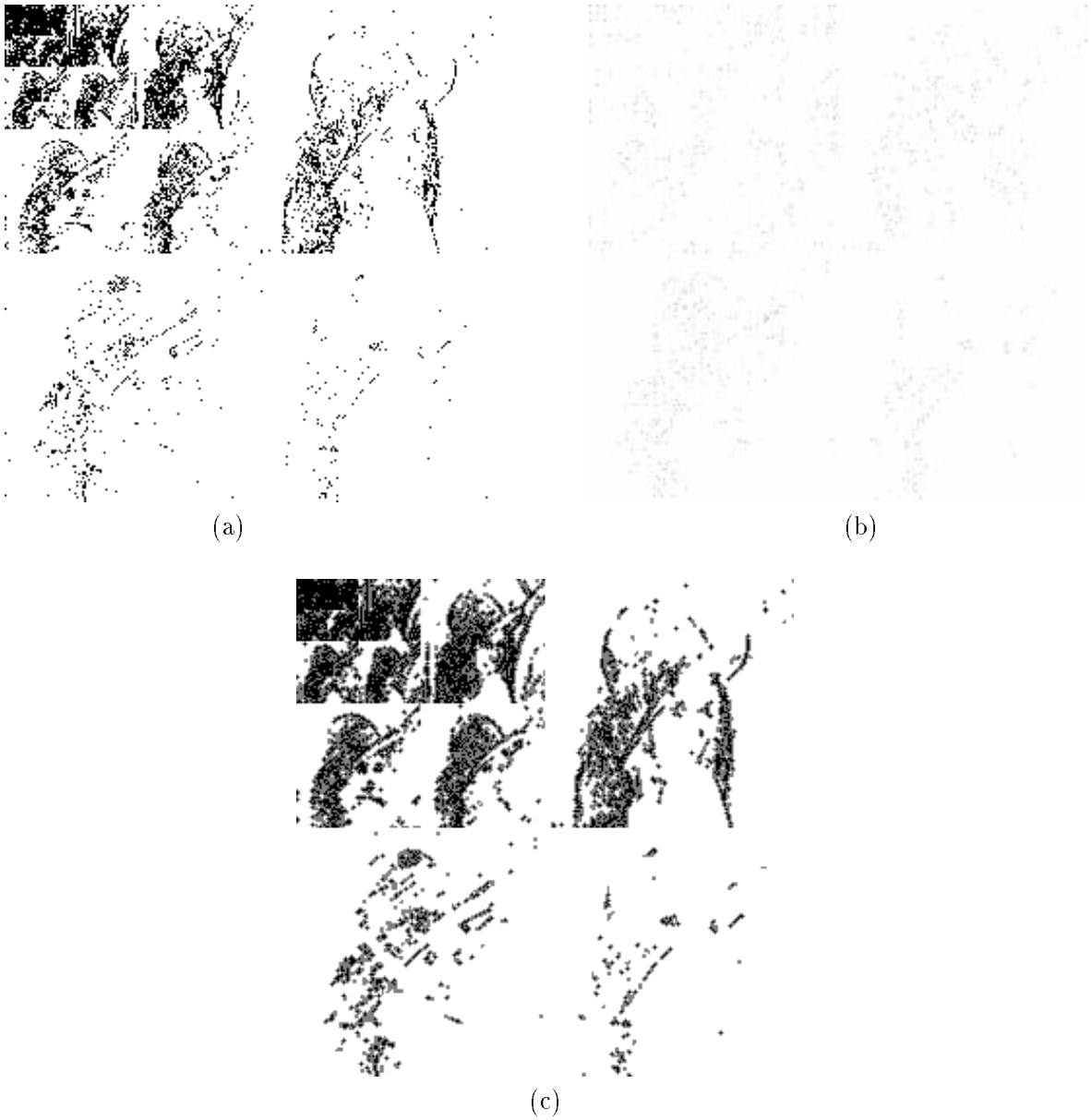


Figure 2: Significance map for six-scale wavelet decomposition, $q = 11$. (a) Significance map after quantization: White pixels denote insignificant coefficients and black pixels denote significant coefficients. (b) The marker image (after removing clusters having only one significant coefficient) shows the seed positions for the 689 clusters. Note that only 16 seed position needs to be transmitted explicitly with the use of significance-link. (c) The transmitted significance map. White pixels denote insignificant coefficients that are not encoded. Black and gray pixels denote encoded significant and insignificant wavelet coefficients, respectively.

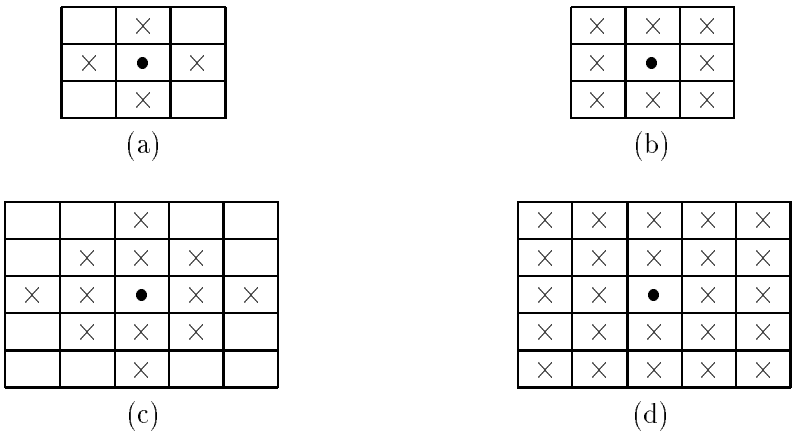


Figure 3: Structuring elements used in conditioned dilation.

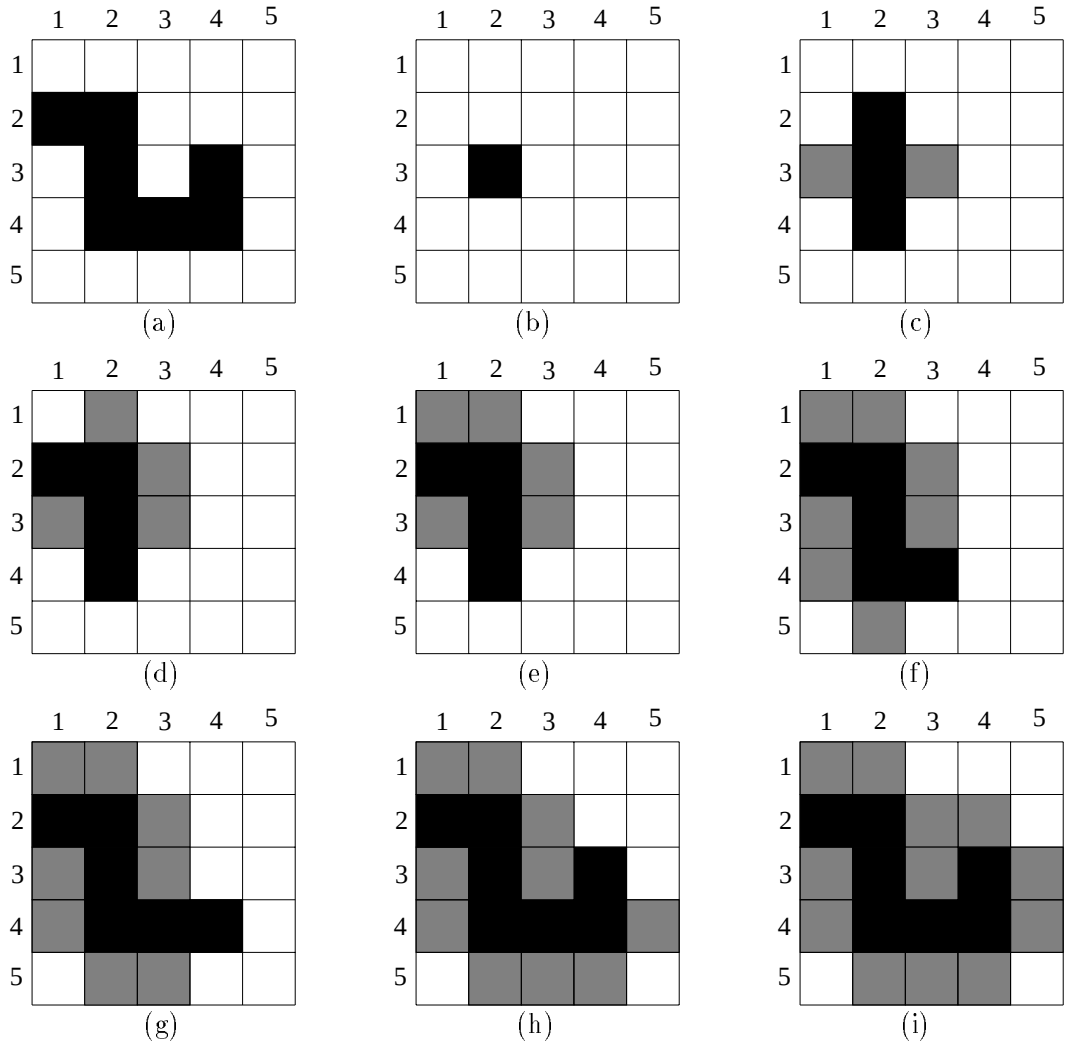


Figure 4: Demonstration of the progressive cluster detection by using conditioned dilation on a simple example. White pixels denote insignificant coefficients that are not coded. Black and gray pixels denote encoded significant (S) and insignificant (I) coefficients, respectively. (a) The significance map and (b) the seed position. (c)-(i) Steps of the algorithm. The final transmitted string is SISIIISISIISISIIII.

Figure	Current	Coded Positions
(c)	(3,2)	(2,2):S, (3,3):I, (4,2):S, (3,1):I
(d)	(2,2)	(1,2):I, (2,3):I, (2,1):S
(e)	(1,2)	(1,1):I
(f)	(4,2)	(4,3):S, (5,2):I, (4,1):I
(g)	(4,3)	(4,4):S, (5,3):I
(h)	(4,4)	(3,4):S, (4,5):I, (5,4):I
(i)	(3,4)	(2,4):I, (3,5):I

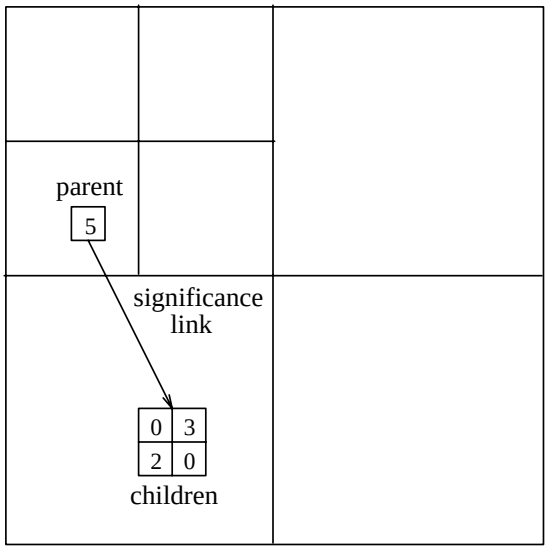


Figure 5: Illustration of significance-link. The values are the magnitudes of quantized coefficients. Nonzero values denote significant coefficients.

(a) (b)

(c) (d)

Figure 6: Coding results for 512×512 "Barbara" image. (a) Original, reconstructed images (b) at 0.5 bpp, PSNR=31.89 dB, (c) at 0.25 bpp, PSNR=28.18 dB, and (d) at 0.125 bpp, PSNR=25.36 dB.

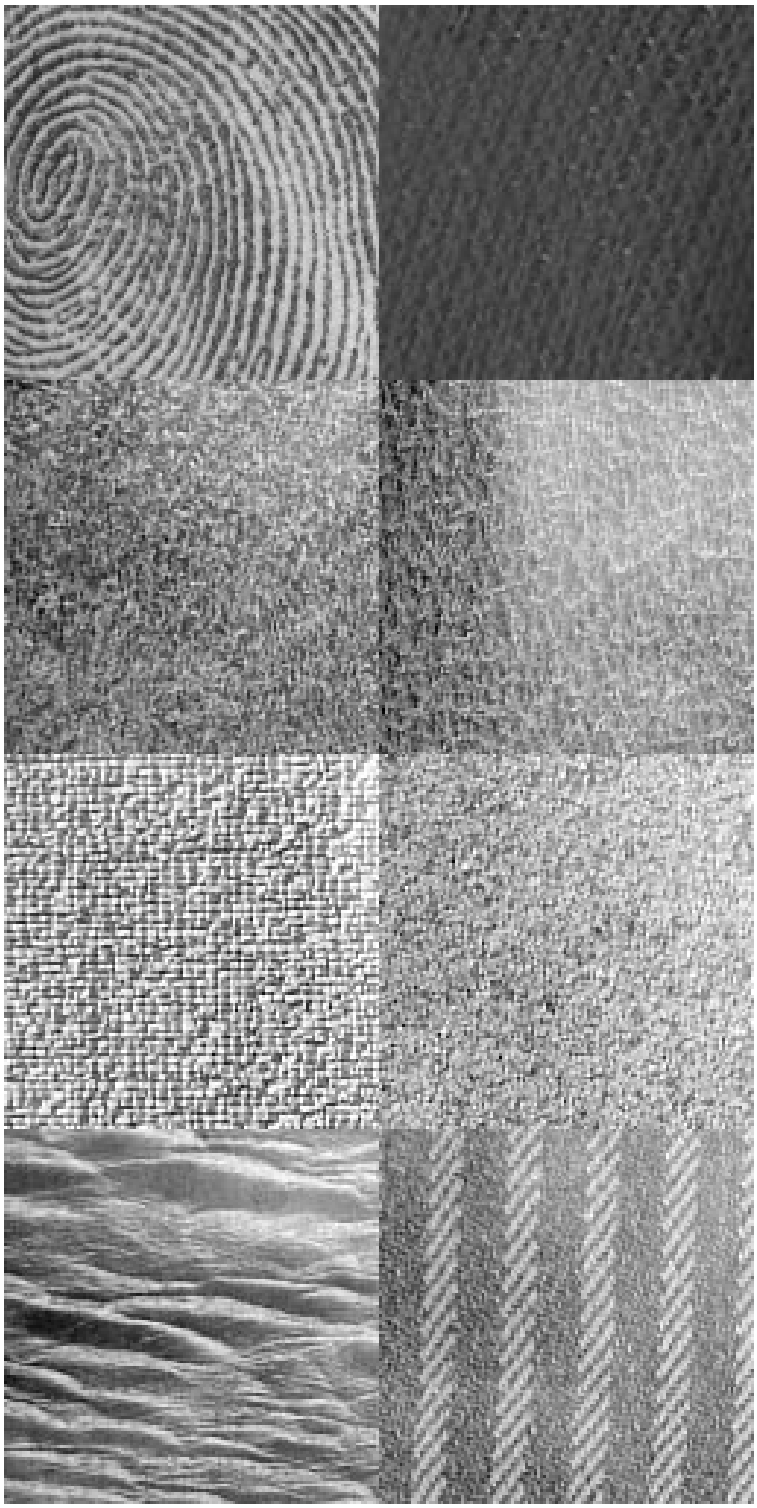
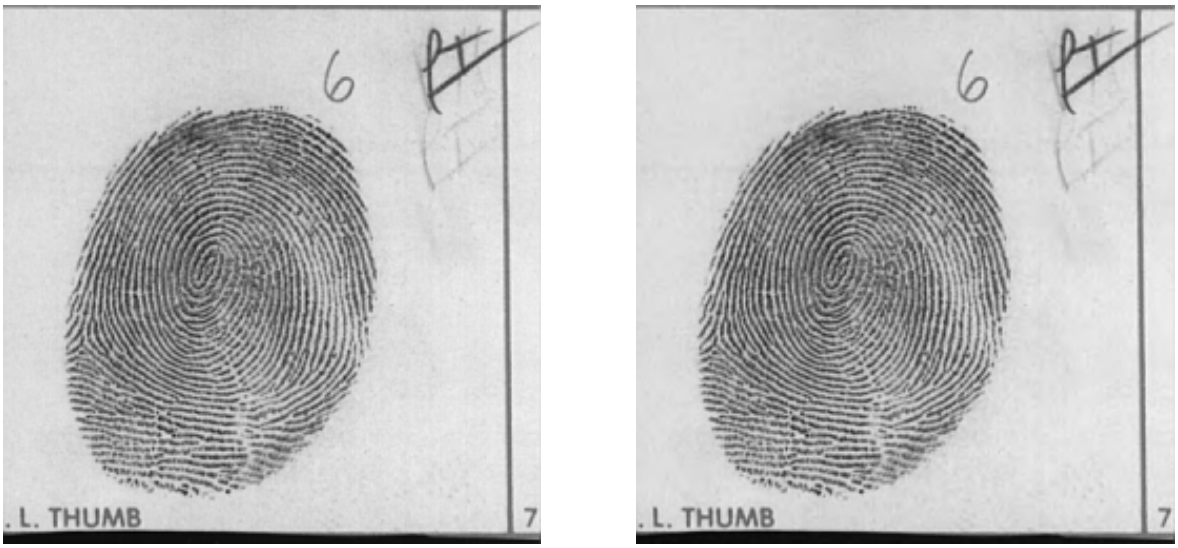


Figure 7: 256×256 texture images. From left to right, top to bottom: "fingerprint," "sweater," "grass," "pig skin," "raffia," "sand," "water" and "wool."



(a) (b)

Figure 8: "Fingerprint" images. (a) Original image. (b) Reconstructed image by SLCCA at 0.444 bpp, PSNR=35.65 dB.