

Wavelet-Based Space-Frequency Compression of Ultrasound Images

Ed Chiu, Jacques Vaisey, and M. Stella Atkins

Abstract—This paper describes the compression of grayscale medical ultrasound images using a recent compression technique, i.e., space-frequency segmentation (SFS). This method finds the rate-distortion optimal representation of an image from a large set of possible space-frequency partitions and quantizer combinations and is especially effective when the images to code are statistically inhomogeneous, which is the case for medical ultrasound images. We implemented a compression application based on this method and tested the algorithm on representative ultrasound images. The result is an effective technique that performs better than a leading wavelet-transform coding algorithm, i.e., set partitioning in hierarchical trees (SPIHT), using standard objective distortion measures. To determine the subjective qualitative performance, an expert viewer study was run by presenting ultrasound radiologists with images compressed using both SFS and SPIHT. The results confirmed the objective performance rankings. Finally, the performance sensitivity of the space-frequency codec is shown with respect to several parameters, and the characteristic space-frequency partitions found for ultrasound images are discussed.

Index Terms—Space-frequency segmentation, ultrasound image compression, wavelet packets.

I. INTRODUCTION

ULTRASOUND is a popular medical imaging modality because it is noninvasive, versatile, there are no known side effects, and the equipment is small and inexpensive relative to other options. Reducing storage requirements and making data access more convenient are two of the motivations for applying compression to ultrasound images, with the retention of diagnostic information being critical.

A distinguishing feature of ultrasound images is the oriented “speckle texture,” which is a consequence of the physics underlying the data acquisition. The speckle energy is typically concentrated in certain spectral regions due to the orientation of the speckle pattern. Depending on the context and application, speckle in medical images can be viewed as signal or noise. For example, speckle can be used to characterize tissue [1], [2] or it can mask diagnostically relevant features [3], [4]. When it is not known *a priori* whether speckle should be viewed as signal or noise in a specific application, the compression algorithm needs to preserve it. Speckle preservation is also important for aesthetic reasons since radiologists are accustomed to it and notice-

Manuscript received May 31, 2000; revised October 18, 2000. This work was supported by the Canadian Natural Sciences and Engineering Research Council.

E. Chiu was with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6. He is now with West Bay Semiconductor, Vancouver, BC, Canada V6H 4C1.

J. Vaisey is with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6.

M. S. Atkins is with the School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6 (e-mail: stella@cs.sfu.ca).

Publisher Item Identifier S 1089-7771(01)07412-X.



(a)



(b)

Fig. 1. Sample ultrasound images. (a) Normal liver U1. (b) Liver lesion U7.

able alterations to images should be avoided when possible. The presence of speckle and the desire to preserve it are issues that distinguish the ultrasound image compression problem from the compression of “natural” images such as landscapes.

Another characteristic of ultrasound images is the spatial variation in pixel statistics across an individual image. A typical image consists of an ultrasound-scanned area, which is often nonrectangular, against a passive background, which may contain text and limited graphics (see Fig. 1). The resulting spatial variation in image statistics presents a challenge to coding methods that use a single partition strategy. For example, many modern image compression algorithms, such as zerotree coding [5] and set partitioning in hierarchical trees (SPIHT) [6] are based on the wavelet transform, which partitions the input into frequency bands whose size decreases logarithmically from

high frequencies to low. This decomposition strategy works well when the input images are statistically homogeneous; however, in general, a wider range of partitions, or bases, may be called for. In addition, since medical applications do not tolerate much loss in fidelity, distortion free methods, such as context-based adaptive lossless image coding (CALIC) [7] have been recently adapted to “near-lossless” situations [8] with good results.

In a comparison of standard techniques, Erickson *et al.* [9] used both SPIHT and JPEG to compress magnetic resonance imaging (MRI) and ultrasound images; they concluded that wavelet-based methods such as SPIHT are subjectively superior to JPEG compressed at moderately high bit rates. In a very recent follow-up study on the compression of ultrasound images, Persons *et al.* [10] compared JPEG- and SPIHT-compressed ultrasound images at low compression rates and concluded that for both, JPEG and SPIHT produced diagnostically acceptable images at a 10:1 compression ratio. However, at compression rates above 12:1, the SPIHT approach developed artifacts, such as lack of low-contrast detail, which impacted the diagnostic acceptability. JPEG compressed images were found “not necessarily acceptable” at compression rates above 15:1 due to blocking artifacts. Persons *et al.* also have a good discussion on the types of artifacts found in SPIHT- and JPEG-compressed images, including the role of the text and graphics, which are often placed on ultrasound images.

Due to prior study on evaluation of compression algorithms for ultrasound images at moderate compression levels, we decided to use SPIHT as the basis for our comparisons. JPEG 2000 provides another possible algorithm to test against, but this method is new and software is not commonly available. Furthermore, although JPEG 2000 usually gives superior performance to SPIHT, the difference has often been reported as being small [11]. We are unaware of any studies comparing the performance of SPIHT and JPEG 2000 on ultrasound images.

The purpose of this paper is to outline the algorithm we used to compress ultrasound images, which is based on the relatively new technique of space-frequency segmentation (SFS) [12]. The paper highlights the methods used to choose the parameters of the algorithm and demonstrates its superiority to SPIHT, both in peak signal-to-noise ratio (PSNR), and in subjective image quality measures. Our paper focuses on the compression of scan-converted ultrasound images that have been interpolated from the original image data by the imaging equipment. We focus on the post-scan-conversion case since this is the display domain and we want to avoid the introduction of “uncontrolled” artifacts through the combination of quantization and interpolation. In addition, issues like bit allocation become difficult to deal with in the pre-scan-converted domain since the interpolation process may cause some quantization errors to be magnified more than others. Studies on compression in the pre-scan-converted domain are left for future study.

A. SFS

When used to compress ultrasound images, SPIHT produces distinctive artifacts. Specifically, speckle is blurred at lower bit rates, suggesting that a different decomposition topology may be better suited to speckly images. Also, since contrast is key

in ultrasound images, bit allocation by coefficient magnitude, as done in SPIHT, is not the best choice. The resulting artifacts are particularly visible in areas of low-contrast detail. To avoid the above problems, we have implemented an image-adaptive scheme that selects the best filter topology. By viewing image decompositions as projections onto sets of basis functions, the core problem is to choose an appropriate basis for each individual image. The drawbacks of an adaptive scheme include the computational cost of finding the best basis and the need to send a description of the basis as side information. The algorithm for finding the optimum representation must have a meaningful criterion for choosing the best basis, and must limit the choice of basis to a useful, but finite, set of possibilities.

SFS starts with a large set of bases, an associated set of quantizers, and then uses a fast tree-pruning algorithm to select the best combination according to a given rate-distortion criterion. As the name suggests, the bases are defined using a hierarchy of space and frequency partitions: space partitioning refers to the division of an image into spatial quadrants, while frequency partitioning refers to the division of an image into four frequency subbands. The resulting subimages can each be successively partitioned in space or frequency, up to some maximum depth, so that a decomposition tree is built. Allowing both space or frequency partitions on any subimage permits more basis choices than previous related methods, such as single- and double-tree wavelet packets [13], which provide only frequency partitions. This additional flexibility makes the resulting compression algorithm appropriate for coding unique image types, such as ultrasound, which are distinct from the more usual “natural” images.

II. CODEC DETAILS

In this section, we describe the SFS codec and present some of the tradeoffs involved in making the parameter choices. In order to evaluate choices, we used two standard “natural” images, i.e., “Lena” and “Barb,” as well as two ultrasound images of size 640×480 pixels. These ultrasound images are labeled U1 and U7 (see Fig. 1) and consist of an ultrasound-scanned area against dark backgrounds, with some light colored text in the background area describing the exam conditions: a bar of gray shades provides a visual indication of the grayscale range. The ultrasound area of the images is wedge shaped, implying that curved linear-array transducers have been used to capture both images. Abdominal exams are a common application for diagnostic ultrasound, and both the test images are of the liver. The images are quite speckly, which is typical for liver images, but it is not clear whether or not the speckle is diagnostically significant. (See [14] for an example of diagnostically significant speckle in liver images.) The text portion contains a small section of inverse video. We use both subjective and objective measures of quality in our evaluations. Although the diagnostic quality of the compressed images is what matters most, i.e., subjective evaluations by radiologists, it is also very useful to use numerical measures that are unbiased and easily replicated. The most common objective quality measure in image coding is the PSNR, which is defined as

$$\text{PSNR} = 10 \cdot \log_{10} \left(M^2 / E \left[(x - \hat{x})^2 \right] \right)$$

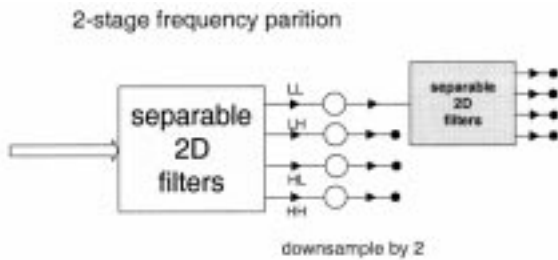


Fig. 2. Subband filtering.

where M is the peak pixel amplitude, x represents a pixel amplitude in the original image, and \hat{x} is the corresponding pixel in the reconstructed compressed image.

The ultrasound test images were chosen to represent a range of compression “difficulty,” while still being typical. The harder case, i.e., U1, is an image of a normal liver captured at 3.5 MHz and shows high contrast and significant fine detail, particularly in the upper central area. There is also low-resolution detail present in the low-contrast dark areas across the center of the image. An unusual feature in U1 is the icon in the upper left-hand side indicating the organ being scanned. The easy case to compress, i.e., U7, was captured using a 5-MHz transducer and shows a liver lesion. At the time of capture, the contrast was adjusted to highlight the lesion so the overall image is quite dark and low contrast. The ultrasound area of U1 occupies a larger proportion of the total image than that in U7. A radiologist commented that U1 is quite typical in terms of quality and content and contains significant textural detail, but has somewhat higher contrast than usual. U7 was evaluated to be a poor quality image without a lot of detail, but since the loss of detail results from the intentional highlighting of the lesion, U7 is not atypical.

A. Subband Decomposition

SFS relies on the concept of subbands, which provide the mechanism for partitioning the image. For example, frequency partitions are implemented by filtering an image region using separable two-dimensional (2-D) filters (one high pass and the other low pass in nature) and then subsampling the four outputs by a factor of two in each dimension. As long as the filters are appropriately designed, then there is no loss associated with this operation. Subband principles have been applied to image coding for many years [15] and the general idea is shown in Fig. 2, where 2-D separable filters are used to form the different subbands LL, LH, HL, and HH, and the labels specify whether a high- or low-pass filter is used first horizontally and then vertically: e.g., LH indicates that the low- and high-pass filters are applied horizontally and vertically, respectively. Fig. 3 shows a single-stage frequency subband decomposition of an ultrasound-only region of a typical liver image: the labels indicate the type of filtering.

In general, subbands are classified into three groups according to the distribution of their coefficients, i.e., high-pass bands, low-pass bands, and space-only bands. The codec treats subbands differently according to their classification. Low-pass subbands are those that have been generated without any

high-pass filter applications and look like subsampled versions of the original image prior to filtering, but with increased dynamic range. The coefficient distribution has similar shape to that of the unfiltered image, but repeated low-pass filtering followed by subsampling concentrates the energy into fewer and fewer coefficients, increasing the range of the coefficient values. High-pass subbands include all those that have been filtered with the high-pass analysis filter at any point in their decomposition hierarchy. High-pass subbands have less energy, and have a large number of zero-valued coefficients. Space-only subbands are partitioned only in space; they are not filtered or subsampled at all. Coefficients in space-only subbands retain the characteristics of a rectangular portion of the original image. The partition gamut for the one-dimensional (1-D) case is shown as a sample 1-D signal decomposition in Fig. 4, where the decomposition can stop at any level as long as the required criteria are met. The wavelet-packet case is shown in the shaded arrows, and data blocks are split into two at each step, using a time or a frequency split.

The subband energy distribution has a large effect on the compression performance and is significantly different between “natural” and ultrasound images. In addition, a controlled capture environment makes the statistical properties of ultrasound images much more predictable than those of natural images. As an example, we consider the case of “Ub” (shown in Fig. 3) and the standard natural image “Lena,” which both have 98% of their energy concentrated into the LL band (neither image contains sharp edges) and show the high-pass energy distribution in Table I. It is clear that this distribution is much more skewed for “Ub” than for “Lena.” Indeed, the only interesting features in “Ub”’s subbands are in the LH subband, a result explained by the elongated horizontal shape of the speckle spots, which itself is due to the different transducer resolution in the lateral and axial directions. Ultrasound images are often oriented such that the direction of insonification is vertical so that the main axis of the speckle spot is in the horizontal direction. We thus anticipate that improved compression performance is possible by adapting the compression algorithm to the unusual characteristics of ultrasound images.

Note that there are many possible filter sets to choose from when performing subband frequency decompositions. We experimentally evaluated several filter sets [16], but settled on the standard 9–7 set of Antonini *et al.* [17] due to its overall good performance. This filter set is commonly used throughout the wavelet compression literature.

B. Quantization and Entropy Coding

As discussed in Section I-B, SFS starts with a family of bases and a set of quantizers for each subimage. The bases set is defined to be all possible space and frequency partitions down to a maximum decomposition depth, which is set *a priori*. This maximum-depth parameter controls the minimum subimage size and serves to balance the complexity and performance. A fast tree-pruning algorithm [13] is then used to select the best combination according to a given rate-distortion criterion.

1) *Quantization:* At each node of its tree, SFS quantizes the corresponding subimage with *each* available quantizer and then entropy codes the result. Entropy coding is a lossless technique

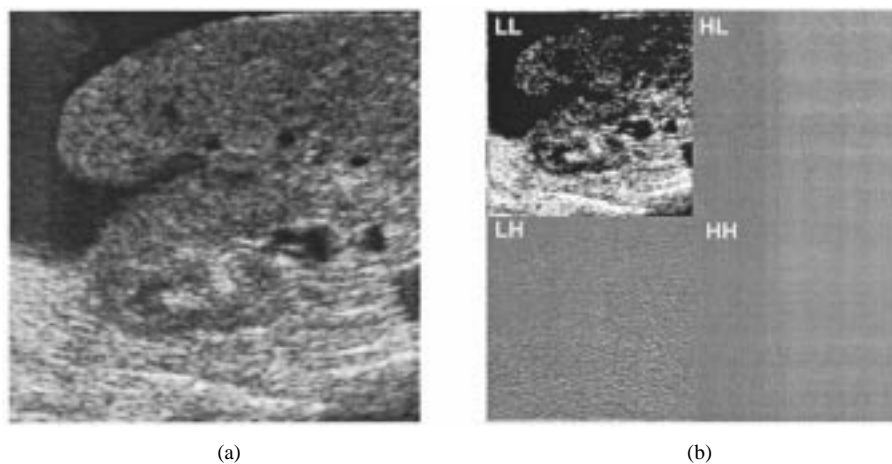


Fig. 3. “Ub.” (a) Ultrasound-only liver image. (b) Subband decomposition.

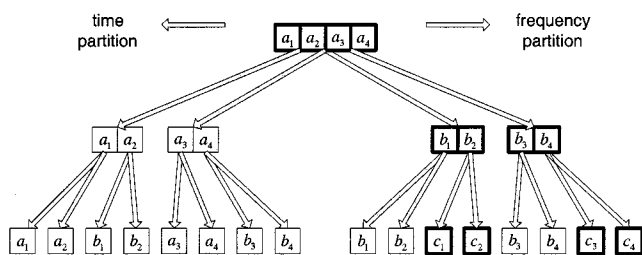


Fig. 4. Space-frequency decomposition.

TABLE I
SUBBAND ENERGY DISTRIBUTION

image	subband	HP energy
Lena	LH	22 %
	HL	64 %
	HH	14 %
Ub	LH	88 %
	HL	10 %
	HH	2 %

that attempts to find the most efficient representation (on average) for the quantizer output. Typical methods for entropy coding [18] are Huffman coding and arithmetic coding, with the latter generally providing better performance. In this paper, we consider the use of two entropy-coding methods, i.e., adaptive arithmetic coding, which is appropriate when the source statistics are initially unknown, and stack-run coding [19], which combines arithmetic and run-length coding and exploits large runs of “zero” coefficients, as is common in many subimages.

The design of the quantizer sets is an important parameter in determining the performance of the codec. All quantizer sets are comprised of 64-level symmetric uniform scalar quantizers, with the step sizes varying according to the subband type. The performance of uniform scalar quantizers has been shown to be similar to that of the best possible scalar quantizer when used in conjunction with entropy coding and assuming independent coefficients [20]. Note that the size of the quantizer sets has a linear

effect on the computational complexity, as well as a small effect on the side information, since the quantizer choice must be sent in the compressed bit stream. We experimentally found that a quantizer set size of 16 gave a good tradeoff between computational complexity and side information to send, and increased coding effectiveness. Herley *et al.* [12] also use sets of size 16. Also, prior to quantization, the original image is preprocessed by subtracting an offset value in order to roughly center the coefficient distribution around zero. For simplicity and to avoid side information, this offset is derived from the precision. For example, an 8-b image with a possible pixel range from 0 to 255 is shifted to have a pixel range from -128 to $+127$.

Our method for finding good step sizes for the quantizer sets starts with a large spread in step size and then refines this range by observing histograms of quantizer usage: ideally, each quantizer should be used a “reasonable” amount. Experiments were also conducted in varying the step-size differences between the set members linearly, or exponentially at a constant rate; however, virtually no difference was found between the two and a linear pattern was used for simplicity. For some subband groups, we found it effective to include a single coarse quantizer (with step-size 128) outside of the linear variation in order to catch outliers, so only 15 out of the 16 total quantizers for these groups have a linear step-size variation.

2) *Entropy Coding*: There is a tradeoff between efficient entropy coding and efficient quantization of SFS images since the “block size” demanded by each may be different. In an ideal world, both the decoder and encoder have sufficient knowledge of the source statistics to build an effective entropy coder; however, in a real implementation, this information has to be communicated. One possibility is to make an initial pass through each subband to estimate statistics, which can then be sent as side information. Unfortunately, this side information requires many bits when the number of subbands is large (as in SFS). An alternative procedure makes the encoder and decoder adaptive so that statistics are estimated “on the fly” using previously coded data. Since this information is available at both ends of the channel, no side information need be sent. Problems still exist, however, since the performance is limited by the accuracy of the estimate and the quality of the estimates depend on both the alphabet size and sequence length. A small alphabet combined

TABLE II
SFS AND SUBIMAGE SIZE

image	scheme	Sub-image size	Encoding time	rate	PSNR
Lena	entropy	Min	0:09	0.4983	36.87
		Opt	0:22	0.4983	36.92
	arithmetic	Min	0:42	0.4995	35.77
		Opt	1:24	0.4972	36.29
Barb	entropy	Min	0:09	0.5018	32.47
		Opt	0:22	0.4997	32.48
	arithmetic	Min	0:42	0.5011	30.99
		Opt	1:24	0.4975	31.67

with a long sequence means that the coder acquires a good estimate of source statistics relatively quickly, and produces an efficient code. On the other hand, the statistics of small subbands, or those with large alphabets, are not estimated well and the entropy coder often performs poorly. Herley [12] avoided this problem by using the zeroth-order entropy to estimate the rate, which is optimistic (and, in fact, results in different partitions); however, this selection also results in a significant reduction in computational complexity—approximately a factor of eight at a typical cost of about 0.3–0.5 dB in PSNR.

An adaptive entropy-coding scheme must balance coding efficiency against simplicity and fast computation. Arithmetic coding is a good choice here and when the alphabet is relatively small and the sequence to be coded is relatively large, an arithmetic coder operating on individual coefficients can achieve a rate very close to the zeroth-order entropy [18]. A coder that exploits correlation between coefficients can do better than the zeroth-order entropy, and the stack-run method [19] does this by combining run-length coding with adaptive arithmetic coding on a small alphabet. To work well, however, this technique needs to see large runs of zero coefficients, implying large subbands.

In determining which entropy coding technique to choose, we experimented with both arithmetic and stack-run coding for the high-frequency bands. The space-only and low-frequency subimages are not appropriate for stack-run coding since they lack the necessary runs of zeros. Our evaluations were made at a target rate of 0.5 bits per pixel (bpp) and stack-run coding was found to perform slightly better than arithmetic coding. In general, the larger the subimage, the more efficient stack-run coding is over arithmetic coding; however, the larger subimages required for coding efficiency in the stack-run case limit the freedom to use many quantizers (since we only use one quantizer for each subimage). Small space partitioned subbands in the arithmetic coding case suggest the use of many different quantizers, but the better quantization choices are balanced against the use of less efficient entropy coding. The power of the SFS algorithm is evident since it applies the strengths of either of the two different entropy-coding schemes to produce quite similar PSNR performance. The two schemes do not differ much in computation time since the extra step of forming stacks and runs followed by arithmetic coding a small alphabet

balances out against the larger alphabet used by arithmetic coding. Since there is not much difference in computational cost, stack-run coding is preferred since, in theory, it can do better than the zeroth-order entropy, even though, in practice, the improvement over arithmetic coding is not as large as hoped for. Note that the relative improvement with stack-run coding was much higher for “natural images.”

C. Complexity Reduction

In its pure form, SFS is a computationally intensive technique since it does a full search using all available bases and quantizers. Nonetheless, it is possible to significantly reduce the computation load at a modest performance cost. We also consider the possibility of using a fixed *a priori* segmentation (which results in a very fast compression) in Section IV-E.

Computation of the rate-distortion costs for each possible partition forms the critical path in the SFS algorithm. Each node in the decomposition tree (see Fig. 4) corresponds to a subimage and decisions must be made whether to partition in space or frequency, or whether to stop partitioning; however, in an ideal world, space partitions do not change the coefficients to quantize and, as long as the quantizer sets do not change, extra space partitions can always be used to descend to the minimum-sized subband without cost. The result is that the rate-distortion cost associated with the “stop partitioning” choice can be avoided at only a small increase in side information (each subimage contains its own quantizer descriptors). As discussed above, however, practical coders have problems dealing with small subbands and there is a penalty associated with making bands smaller than necessary. To evaluate the tradeoff, we quantized “Lena” and “Barb” using both the optimum decomposition and that obtained when always using minimum-sized subimages. Table II reports the experimental results and it is seen that using minimum-size subbands (in this case, 16×16) reduces the computation by a factor of around 2.5 at a “reasonable” cost. Herley *et al.* [12] restrict the subimages to the minimum size, but because they use zeroth-order entropy to estimate the rate instead of actually coding the subimages, they observed no penalty and their results are somewhat misleading.

The computation time depends on the maximum decomposition depth, thus, a convenient way to speed up the algorithm is to

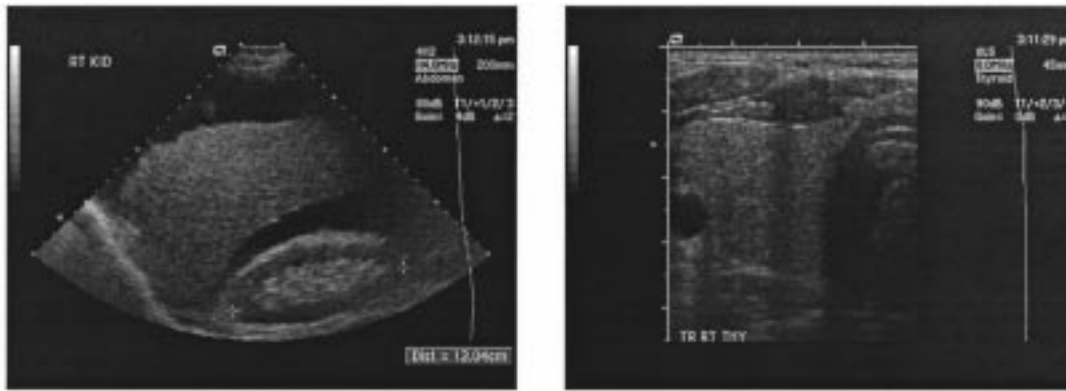


Fig. 5. Images for the subjective quality evaluation.

shrink this number. For example, decreasing the depth from five to four to three reduces the encoding time by approximately 40% and 60%, respectively. We performed experiments on images U1 and U7 with decomposition depths varying from three to five, at a target bit rate of 0.5 bpp and the PSNR results were found to improve by between 0.5–1.25 dB for each unit increase in depth. We set this parameter to five for the remainder of our experiments.

Since the computational load of the search is linearly dependent on the number of quantizers in each set, we varied this parameter from 4 to 16 and observed the PSNR results when coding U1 and U7. The results showed that there is essentially zero performance penalty associated with reducing the number of quantizers from 16 to 8; however, with fewer quantizers to choose from, the performance of the codec becomes more sensitive to individual image statistics. As long as speed is not critical, it is thus useful to keep this number as high as possible and we used 16 quantizers in each set in the following.

III. EXPERIMENTAL DESIGN

In the experiments described below, the SFS encoder and decoder were implemented as separate programs that communicated only through a file containing the compressed image representation. This configuration means that all coded rates reported for the implemented SFS codec are reliable, being obtained from the size of the representation file. This file did not include the image width and height, but did include the side information specifying the optimum partition and the quantizer for each subband. Our testing procedure involved objective quality measures, specifically the PSNR, as well as a subjective evaluation by radiologists. These experiments used slightly different target rates, as described below. Note that the compression rates in the tests are not always identical since the SFS algorithm produces the best partition/quantizer combination that it can at a rate below the target. For this reason, it is often necessary to iterate the target to get close to the desired rate.

An appropriate target rate had to be determined prior to subjective testing. After a number of experiments, this target was set to 0.4 bpp, a rate that resulted in coded ultrasound with visible artifacts while maintaining a relatively high subjective quality. A radiologist who examined U1 and U7 compressed at 0.4 bpp using SFS confirmed the choice of target bit rate, commenting that artifacts in U1 are present, but image quality is

still quite good, while artifacts in U7 are hard to detect, and the slight edge sharpening around the lesion in the compressed version of U7 actually makes the compressed version preferable to the original. As described above, the maximum decomposition depth is five and there is a choice between 16 different 64-level quantizers for each subband: i.e., the high-pass, low-pass, and space-only subbands choose from a different quantizer sets. Details on the exact quantizer sets used in the experiments can be found in [16].

We did not have the resources to perform a full formal study of compressed medical image quality; however, we were able to conduct a limited set of subjective quality tests using two volunteer radiologists from the Vancouver Hospital and Health Sciences Centre (VHHSC), Vancouver, BC, Canada. The mandate of this study was to compare the subjective quality with respect to medically relevant features of ultrasound images compressed with SFS and SPIHT at a single bit rate. The study was conducted in the radiology viewing room using monitors that are normally used for ultrasound images, under typical lighting conditions. Although such a small study is not sufficiently large to draw general conclusions regarding the diagnostic and perceptual quality of SFS as compared with SPIHT, it does show that there is a quality difference in some images and, thus, motivates future and more extensive work in this area. The 12 images in the test set were chosen with the help of a radiologist to cover a selection that is representative of what radiologists encounter normally, and which include diagnostically interesting features. Since the VHHSC Radiology Department conducts a large number of abdominal exams, the images were dominated by images of the liver, kidney, gall bladder, and pancreas. Non-abdominal images in the test set consisted of two thyroid images. Most of the abdominal images were captured at 4 MHz using harmonic imaging, in which signal processing using the second harmonic produces a slightly sharper image. The rest of the abdominal images were captured at 3.5 MHz. Curved linear- and phased-array transducers were used for the abdominal images. The thyroid images were quite different from the other images because they were captured at 8 MHz using linear-array transducers and, consequently, the ultrasound portion was rectangular in shape. All images in the test set were grayscale with a resolution of 8 b. The PSNR of all the compressed test images ranged from 35.1 to 41 dB. Two of these test images are shown in Fig. 5 (the full set is available in [16]).

The test methodology was the two-alternative forced-choice technique. Subjects were shown a pair of images and selected the one with the best subjective quality in terms of medically relevant features. Each pair had a common uncompressed image and, thus, there were three possible permutations, i.e., space-frequency versus SPIHT, uncompressed versus space-frequency, and uncompressed versus SPIHT. Twelve pairs of images were used, and since the goal of the test was to compare SFS against SPIHT, six of the 12 pairs were space frequency versus SPIHT, while three pairs compared uncompressed versus space frequency and three pairs compared uncompressed versus SPIHT. Within the three possible permutations there is a mix of image quality (as measured by PSNR) to minimize the bias toward either space frequency or SPIHT. To limit learning effects, presentation permutations were interleaved in the order: 1) space frequency versus SPIHT; 2) uncompressed versus SPIHT; 3) space frequency versus SPIHT; and 4) uncompressed versus space frequency. Images were presented side by side, but the higher quality image in terms of PSNR was randomly distributed to the left- or right-hand side. The subjects were not told whether the images were compressed or which compression methods were used. A numerical quality scale was not used since there was not enough variation in quality at a single bit rate to justify this method; however, the radiologists were encouraged to record any relevant comments about quality.

IV. RESULTS AND DISCUSSION

A. Space-Frequency Partitions

It is interesting to observe the partition found by the SFS algorithm. A dramatic way to show these partitions is via an image of subimages, and we adopt this approach in Fig. 6. These images show the final set subimages used by the coder when operating at a target rate of 0.5 bpp: black and white lines are used to denote frequency and space partitions, respectively. The partitions appear as successive splitting into quadrants, with the length of the lines showing the size of subimage that was split. For example, the first partition divides the image into four subimages and, in the case of U1 and U7, the full-length horizontal and vertical white lines indicate an initial spatial domain split. The shape of the ultrasound-scanned region drives the early partitions since the background and test areas are best coded spatially. Since space partitions are limited to binary divisions of the width and height, it is difficult to fit the wedge-shaped ultrasound area well. Multiple space partitions early in the decomposition hierarchy limit the depth of subsequent frequency partitions on the ultrasound-scanned area, and produce small subbands that limit coding efficiency after only a few frequency partitions.

When the space partitions isolate an ultrasound-scanned area, frequency decompositions dominate, and since the depth is limited by the early space partitions, the frequency partitions proceed to the maximum depth in order to accomplish energy compaction. Note that, after the first frequency partition, the LH subband is often further partitioned in frequency, but the HL and HH bands are not. As shown earlier in Fig. 3, the high-pass speckle energy is concentrated in the LH band when the image is

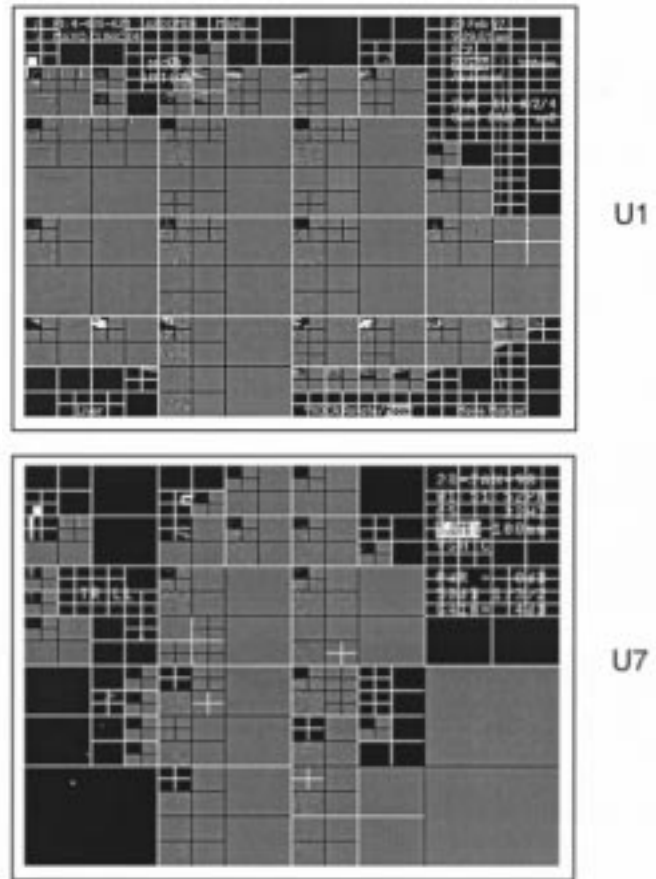


Fig. 6. Segmented ultrasound images.

oriented such that speckle spots are elongated in the horizontal direction; this is also the case in Fig. 6, which explains why additional decompositions of the LH band produce additional energy compaction.

The SFS decompositions show that the best way to compress ultrasound images is to treat the background and ultrasound scanned areas differently. However, this is not a simple task because of the variety of different objects that can appear in the background (text, reverse video text, graphic objects), and the unusual shape of the ultrasound scanned region. SFS has limitations in dealing efficiently with both background and ultrasound areas simultaneously, but because it finds the optimal representation without making any assumptions about the underlying image, and because it incorporates space partitions that allow it to deal with different regions with different characteristics, the algorithm performs relatively well.

B. Objective Image Quality and Comparisons with SPIHT

The PSNR versus bit-rate results are shown in Fig. 7 and it is evident that SFS outperforms SPIHT over a wide range of bit rates. In addition to its relatively poorer PSNR performance, SPIHT also suffers from a noticeable degradation in the subjective quality even at high PSNR values. Since it allocates more bits to large magnitude coefficients in low-frequency bands, SPIHT seems to have trouble with the speckle and other fine details in ultrasound images. Focusing in on a specific bit rate, Table III shows the PSNR at 0.5 bpp, along with the required

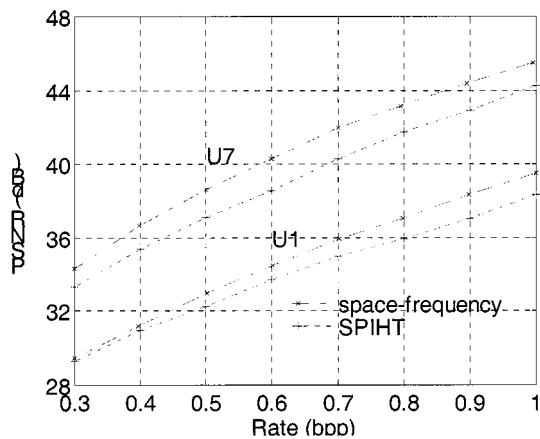


Fig. 7. Compression performance at different bit rates.

TABLE III
PSNRS AT 0.5 bpp

Image	Codec	PSNR (dB)	Coding time (mins:secs)
U1	SFS	32.99	2:15.00
	SPIHT	32.26	0:00.20
U7	SFS	38.60	1:52.00
	SPIHT	37.10	0:00.20

execution time on a Pentium II 400-MHz computer. It is clear that SFS takes much longer than SPIHT to execute, although this time could be substantially reduced through optimizing the code, and the use of a smaller quantizer set and/or a common partition. Also, the SFS decoder is very fast since the only computationally intensive component is the inverse transform.

Comparing the space-frequency compressed version of U1 to the original, we find that compression has resulted in a minor blurring of fine high-contrast textural detail, as well as some loss of low-contrast detail in the dark areas across the center of the image. An interesting artifact that is present in the SFS compressed U1, but not the SPIHT compressed U1, is the faintly visible boundary between space-partitioned regions. The boundary is most noticeable for space partitions that occur early in the decomposition hierarchy. This boundary artifact is analogous to the blocking effect in transform-coded images and is only produced by space partitions. An informal comparison between the SFS and SPIHT versions of U1 shows that SPIHT preserves less detail and adds more artifacts than SFS.

In order to highlight the above effects, we zoom in on specific regions of U1: the improvement in subjective quality of SFS over SPIHT is less noticeable for U7 than U1 since the characteristics of U7 make it easier to compress. Fig. 8 shows a “typical” detailed region, and it is seen that the SFS image has significantly less blurring than the SPIHT one. Also note the vertical space partition boundary visible in the SFS image about 1/3 of the way across the zoomed area. An area with low-contrast detail is shown in Fig. 9 and the extra blur caused by SPIHT is again evident.

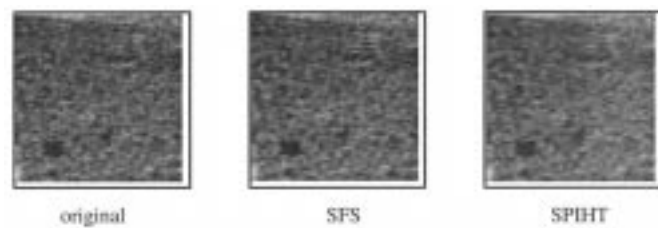


Fig. 8. Compression results in finely detailed speckle regions of U1.

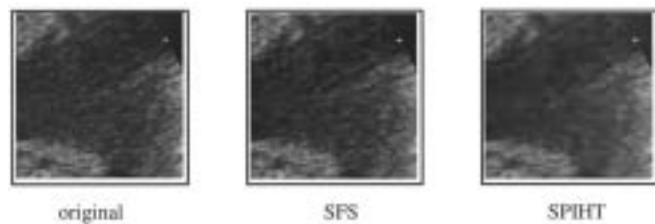


Fig. 9. Compression results in a low-contrast region of U1.

TABLE IV
SUBJECTIVE RANKINGS

Comparison	Category	Winners
SFS vs SPIHT	Space-Frequency	11
	SPIHT	1
	No decision	0
Original vs SFS	Uncompressed	3
	Space-Frequency	1
	No decision	2
Original vs SPIHT	Uncompressed	5
	SPIHT	0
	No decision	1

C. Results of Expert Viewer Assessments

As described in Section III, two radiologists did a series of subjective evaluations and the winners of the 24 comparisons are shown in Table IV. Due to unclear instructions to one of the radiologists, there are three “no-decision” results, where the radiologist cannot choose a winner from the pair of images.

The 12 comparisons between compressed images showed a clear preference for SFS compression over SPIHT, with SFS being preferred 11 times out of 12, while SPIHT won only once. In the one case where SPIHT was preferred, i.e., a thyroid image, the quality of both compressed images was quite high—39.7 dB for space frequency and 38.8 dB for SPIHT. The radiologist who chose the SPIHT-compressed image preferred it because it appeared to have more contrast and sharpness. At high-quality levels where objectionable artifacts are minimal, the low-pass filtering effect of the wavelet transform can sharpen features by reducing high-frequency noise, so that the effects of compression can produce an image that is preferred over the original.

Despite the clear preference for SFS images over those of SPIHT, the distinction in quality between the two compression methods with respect to medically relevant features was not

large at the PSNRs tested. The radiologist's comments indicated that the winner was often chosen based on better sharpness and higher resolution, particularly in speckly texture regions of an image. In a few cases, the SPIHT compressed image lost because visible artifacts degraded medically significant features. At 0.4 bpp, most of the compressed images in the test set had some loss of detail and exhibited some compression artifacts, but the utility of most of the images for diagnostic purposes was rated as not significantly degraded. When the radiologists focused on compression artifacts, both were always able to identify the image in a pair with lower PSNR, but both radiologists concentrated on using medically significant features as the basis of their winner choice, rather than simply choosing the image with the least visible artifacts. Both radiologists commented that all but one of the images that they were shown were suitable for diagnostic use. The one image that was judged unsuitable was a SPIHT compressed image in which severe artifacts and blurring were visible. The unsuitable image was a low-contrast image of a kidney. The bit allocation scheme used in SPIHT, which allocates more bits to larger magnitude coefficients, appears to have trouble with low-contrast images where the coefficients have similar magnitudes.

In the six comparisons between SFS compressed and uncompressed images, the SFS image was the winner once, the uncompressed image was the winner three times, and there were two "no decisions." In the six comparisons between SPIHT compressed and uncompressed images, the uncompressed image was the winner five times, and there was one "no decision." Thyroid images appeared easier to compress than abdominal images, and the "no-decision" case for the SPIHT versus uncompressed comparison involved a thyroid image where the PSNR of the SPIHT compressed image was quite high (41 dB).

D. Compression of Ultrasound-Only Images

To focus more closely on the medically relevant sections of the ultrasound images, experiments were performed on 192×256 pixel regions cropped from five medical ultrasound images with wedge-shaped ultrasound scanned areas.¹ These images, labeled C1–C5, are shown in Fig. 10, and it is seen that they range in contrast from quite high (C1) to quite dark (C4). All images are quite speckly, and C3–C5 also contain features from imaged organs.

The subband images describing the optimum partition for each of the above ultrasound-only images are shown in Fig. 11. Interestingly enough, in each case, the initial low-pass band is recursively partitioned in frequency, as in the wavelet transform [17] (recall that black lines denote a frequency partition and white lines denote a space partition). The lack of spatial partitions in the ultrasound regions suggests that it could be advantageous to segment the image into ultrasound/nonultrasound regions prior to coding, which is exactly what SFS tries to do when partitioning the entire image. Unfortunately, a general segmentation would require the coding of nonrectangular regions (as in [21]) and could leave artifacts near region edges if not done carefully. The frequency distribution in the subimages also shows

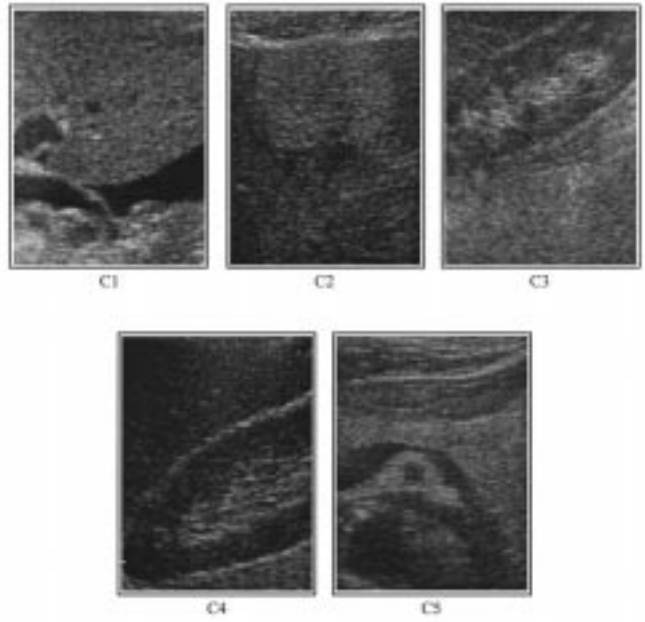


Fig. 10. Ultrasound-only test images.

that ultrasound images are still generally low pass in nature, despite the presence of speckle. All five optimum partitions also include a frequency partition of the LH band at depth one. Partitioning of the LH band is consistent with the subband decomposition experiment in Section IV-A, which showed significant energy in the LH band due to the shape of the speckle spots. We also observe that space partitions are rare, which is likely due to the relative homogeneity of the ultrasound-only images.

E. Encoding Using a Common Partition

The main drawback of the SFS scheme is the large computational load of the algorithm, which is dominated by the effort needed to find the best partition. Since ultrasound images are similar in many ways (especially those that are obtained using similar equipment and which image the same organs), it is reasonable to think that the partitions discovered by the SFS will be reasonably consistent and that a fixed *a priori* partition might perform well. To test this above hypothesis, we compared the partitions shown in Fig. 11 and obtained this "common" partition by comparing the independent choices at each decision point and selecting the winner by a majority vote. The resulting partition is identical to the optimal partitions chosen for images C1, C3, and C5, and is presented in Fig. 12. We have highlighted the octave-band subband decomposition in Fig. 12 to clarify the difference between it and the wavelet-packet-based "common" partition.

The PSNR from the common-partitioning encoding showed that there is little penalty (less than 0.1 dB at 0.5 bpp) in using the common partition, at least when the ultrasound data has been isolated. In both cases considered, the common partition still performs better than SPIHT, particularly in regard to preserving the speckle detail. This result suggests that *a priori* knowledge of the image structure may allow for drastic cuts in the number of partitions to search, resulting in a fast version of the SFS algorithm.

¹C1 is cropped from U1, and C2 is cropped from U7. The original images used to produce C3–C5 were obtained from the Radiology Department, VHHSC.

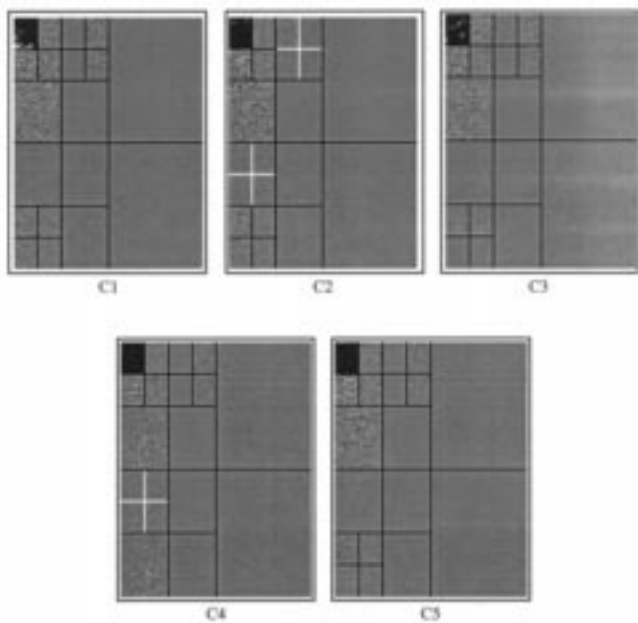


Fig. 11. SFS of ultrasound-only images.

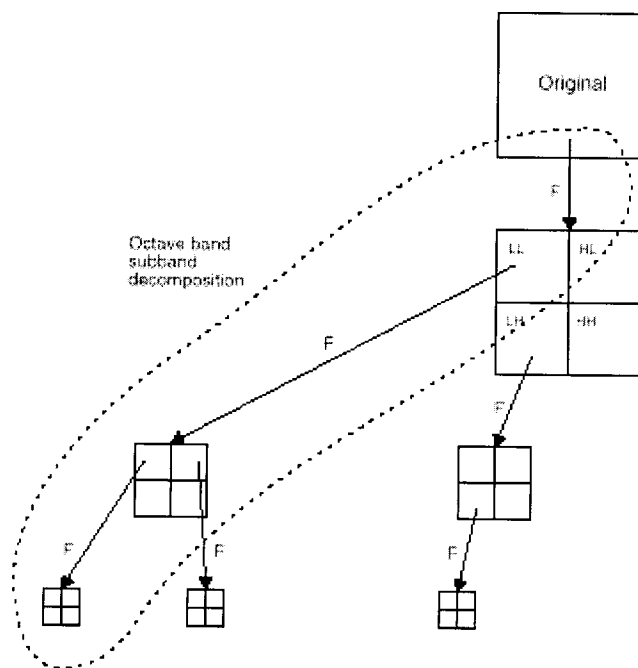


Fig. 12. The "common" partition for ultrasound images.

V. CONCLUSIONS

SFS uses the rate-distortion criterion to find the best representation of an image from a set of bases and quantizers. Since the representation is adapted to suit the image, SFS is a good choice for images with unusual features, such as speckle texture and the unique shape of the scanned region in ultrasound images. The optimum partition for the space-frequency representation indicates that the best way to compress medical ultrasound images is to treat the background and ultrasound scanned areas differently.

Experiments coding ultrasound images with a space-frequency codec produce good results, showing improvement in

terms of both PSNR and subjective quality when compared with wavelet-transform zerotree coding using SPIHT. The improvement in subjective quality is particularly noticeable in regions of fine speckle texture. SFS does a better job of preserving both fine detail and low-contrast detail. It should be noted, however, that a formal study with a large number of radiologists and images is necessary to strengthen the observations and conclusions made in this paper regarding comparisons between SFS and SPIHT. As it stands, our subjective study has been too small to be conclusive.

Although the computational complexity can be reduced significantly by varying the algorithm parameters, the high computational cost remains the primary drawback in using SFS to code ultrasound images. However, it is possible to mitigate this problem by adjusting various codec parameters and perhaps by using a common partition.

ACKNOWLEDGMENT

The authors would like to thank E. Chen, School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, for helping with this paper's data and figures.

REFERENCES

- [1] R. F. Wagner, M. F. Insana, and D. G. Brown, "Unified approach to the detection and classification of speckle texture in diagnostic ultrasound," *Opt. Eng.*, vol. 25, pp. 738–742, 1986.
- [2] M. F. Insana, R. F. Wagner, B. S. Garra, and T. H. Shawker, "Analysis of ultrasound image texture via generalized Rician statistics," *Opt. Eng.*, vol. 25, pp. 743–748, 1986.
- [3] P. M. Shankar, "Speckle reduction in ultrasound B-scans using weighted averaging in spatial companding," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. UFFC-33, pp. 754–758, 1986.
- [4] X. Zeng, E. A. Geiser, A. F. Laine, and D. C. Wilson, "Homomorphic wavelet shrinkage and feature emphasis for speckle reduction and enhancement of echocardiographic images," *Proc. SPIE*, vol. 2710, pp. 658–667, 1996.
- [5] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.
- [6] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. for Video Technol.*, vol. 6, pp. 243–250, June 1996.
- [7] X. Wu and N. Memon, "Context-based adaptive lossless image coding," *IEEE Trans. Commun.*, vol. 45, pp. 437–444, Apr. 1997.
- [8] X. Wu and P. Bao, "L-infinity constrained high-fidelity image compression via adaptive context modeling," *IEEE Trans. Image Processing*, vol. 9, pp. 536–542, Apr. 2000.
- [9] B. J. Erickson, A. Manduca, P. Palisson, K. R. Persons, D. Earnest, and V. Savcenko, "Wavelet compression of medical images," *Radiology*, vol. 206, pp. 599–607, 1998.
- [10] K. R. Persons, P. M. Pailisson, A. Manduca, W. J. Charbonneau, and E. M. James, "Ultrasound grayscale image compression with JPEG and wavelet techniques," *J. Digital Imaging*, vol. 13, pp. 25–32, 2000.
- [11] D. Taubman, E. Ordentlich, M. Weinberger, G. Seroussi, I. Ueno, and F. Ono, "Embedded block coding in JPEG2000," in *IEEE Int. Image Conf.*, vol. II, Vancouver, BC, Canada, 2000, p. 33.
- [12] C. Herley, Z. Xiong, K. Ramchandran, and M. T. Orchard, "Joint space-frequency segmentation using balanced wavelet packet trees for least-cost image representation," *IEEE Trans. Image Processing*, vol. 6, pp. 1213–1230, Sept. 1997.
- [13] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, vol. 2, pp. 160–175, Apr. 1993.
- [14] J. Krasner, S. B. Lo, B. Garra, and S. K. Mun, "Effect of vector quantization on ultrasound tissue characterization," *Proc. SPIE*, vol. 1653, pp. 360–373, 1992.

- [15] J. Woods and S. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1278–1288, 1986.
- [16] E. Chiu, "Lossy compression of medical ultrasound images using space-frequency segmentation," M.A.Sc. thesis, School Eng. Sci., Simon Fraser Univ., Burnaby, BC, Canada, 1999.
- [17] M. Antonini, M. Barlaud, and P. M. I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205–220, Apr. 1992.
- [18] K. Sayood, *Introduction to Data Compression*. San Mateo, CA: Morgan Kaufmann, 1996.
- [19] M. J. Tsai, J. D. Villasenor, and F. Chen, "Stack-run image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 519–521, Oct. 1996.
- [20] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 485–497, 1984.
- [21] T. Sikora, "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments," *Signal Process. Image Commun.*, vol. 7, pp. 381–395, 1995.

Ed Chiu received the B.A.Sc. degree from the University of British Columbia, Vancouver, BC, Canada, in 1988, and the M.A.Sc. degree in engineering science from Simon Fraser University, Burnaby, BC, Canada, in 1999.

He is currently with West Bay Semiconductor, Vancouver, BC, Canada.



Jacques Vaisey received the M.Sc. degree from Queen's University, Kingston, ON, Canada, in 1982, and the Ph.D. degree from the University of California at Santa Barbara, in 1989, both in electrical engineering.

From February 1982 to September 1984, he was with Bell-Northern Research, Ottawa ON, Canada, where he was involved with the design of digital radio systems. From September 1989 to September 1990, he was a Research Associate with INRS-Télécommunications, Montréal, QC, Canada. He is currently an

Associate Professor of engineering science at Simon Fraser University, Burnaby, BC, Canada. His current research interests include video and image coding, quantization theory, image processing, and multidimensional signal processing.



M. Stella Atkins received the M.Phil. degree from the University of Warwick, Warwick, U.K., in 1976, and the Ph.D. degree in computer science from the University of British Columbia, Vancouver, BC, Canada, in 1985.

She is currently a Professor in the School of Computing Science, Simon Fraser University, Burnaby, BC, Canada, and Director of the Simon Fraser University Center for Systems Science. Her research interests include medical image analysis and visualization for diagnosis and therapy, teleradiology, and the

user interface for computer-assisted diagnosis and surgery.