Literature Survey: Compression of Synthetic Aperture Sonar Images

Thomas Higdon

March 24, 2008

Abstract

Many contemporary sonar systems have thousands of elements and sampling rates exceeding 10 kHz. With single-precision floating point values representing data samples, this can mean data rates exceeding 1 Gbps, far exceeding the capabilities of current underwater communication systems while also quickly overwhelming the typical data storage capacity of most underwater vehicles. Image compression techniques have potential for reducing the amount of data required to represent these signals. This paper reviews the current state-of-the-art in wavelet compression techniques and assesses the applicability of these methods to sonar data.

Modern sonar systems generally consist of a large array of signal transmitting and receiving elements. The number of these elements can be in the thousands. Certain sonar applications require relatively high frequencies (often in excess of 100 kHz). The sheer volume of data makes systems like these unwieldy in certain situations. The bandwidth required for the transmission of this raw data is far beyond the capabilities of any current underwater communication system. Furthermore, for remotely or autonomously operated underwater vehicles, the required data bitrate makes storage on physical media for more than a few seconds very difficult.

This project will address the application of image compression techniques to synthetic aperture sonar image data. Raw sonar data may not correlate well with human visual data, so the performance of transform-based techniques will be assessed to determine their usefulness for this type of data. A data set for a sonar system designed at the Applied Research Laboratories at the University of Texas[1] will be used for the investigation of these methods.

One objective of this survey will be to introduce the concepts necessary for the understanding of the image compression techniques that are applicable to sonar. The wavelet transform is integral to the techniques that will be studied, and a brief overview of the subject will be presented. Another objective will be to present applications of the wavelet transform to image compression, including academic journal papers on wavelet zerotrees and Set Partitioning in Hierarchical Trees (SPIHT). The final objective will be to discuss the suitability of these ideas for sonar imagery and advance possible modifications of these techniques to tailor them for sonar data sets.

Transform-based compression techniques usually consist of three steps: Transforming the data, quantizing the coefficients, and lossless compression of the result. There is often more weight given to low-frequency data than high-frequency data, because in typical optical images, more information detectable by the human visual system is stored there. However, in sonar images, this may not be the case. An important characteristic of a sonar image could be the high contrast of a particular edge, and too much compression using a discrete cosine transform technique, for example, could blur an edge like this, causing information loss. Studies using data from synthetic aperture radar have found that the use of the discrete wavelet transform (DWT) to be superior to the DCT for data compression without quality loss [2].

1 The Wavelet Transform

Before discussing compression schemes that are based on the wavelet transform, it is useful to review some of the concepts that underlie its use. Early image compression techniques, like JPEG, are based on the discrete cosine transform (DCT), which is very much like the discrete Fourier transform (DFT). The DCT (and DFT) transform an image from the discrete space domain to the discrete spatial frequency domain. If, for example, a DCT is done across an entire image, the resulting transform coefficients will be less correlated with each other than the pixels in the space domain. However, because the transform is done across the entire image, and different segments of the image may not be correlated with each other, there is more room for decorrelation in the space domain than the DCT by itself offers. JPEG addresses this problem by segmenting the image into 8x8 blocks, depending on the notion that pixels in a small area are usually similar and therefore the information in that block can be represented in fewer bits. Similarity between blocks at all scales is not taken advantage of, and this is a fundamental limitation of Fourier transform-based techniques.

The wavelet transform can be thought of as an extension to the short-time Fourier transform (STFT) or windowed Fourier transform. In 1-D, the STFT windows the signal using a certain window length and takes the Fourier transform of that windowed signal. This windowed transformation is done at each point in time to yield a time-frequency plot called a spectrogram (Figure 1) [3]. The STFT can also be thought of as a filter bank, where each filter corresponds to a frequency point, and the filter is the impulse response of the window. The time points are filtered with each filter in the filter bank across all points in time.

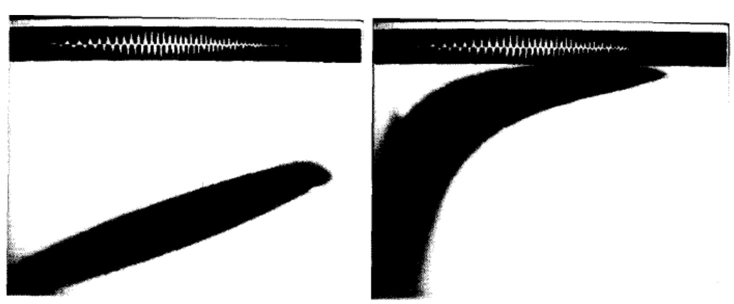

Figure 1: Left: Spectrogram of a sweep signal. Right: Scalogram of a sweep signal.

The STFT is limited by the fact that increasing the resolution in the the time domain involves decreasing the length of the window which causes a corresponding decrease in resolution of the frequency domain. This is because a narrower window indicates a shorter filter, which cannot distinguish smaller differences in frequency. The continuous wavelet transform addresses this problem varying the width of the frequency window according to the frequency being measured. This allows the measurement of arbitrarily small differences in frequency, while at the same time allowing the resolution of arbitrarily small differences in time. This resolution is done on a logarithmic scale. The equivalent to a spectrogram for the wavelet transform is called a scalogram (Figure 1).

We have discussed at a high level a particular interpretation of the continuous wavelet transform, but similar concepts apply to the discrete domain. The typical model of a discrete wavelet transform (DWT) is a cascade of filters, illustrated in Figure 2. Typically $h(n)$ is a low pass filter from $-\pi/2$ to $\pi/2$, and $g(n)$ is the opposite high pass filter. The cascade typically takes place until the signal has been downsampled to a single sample.

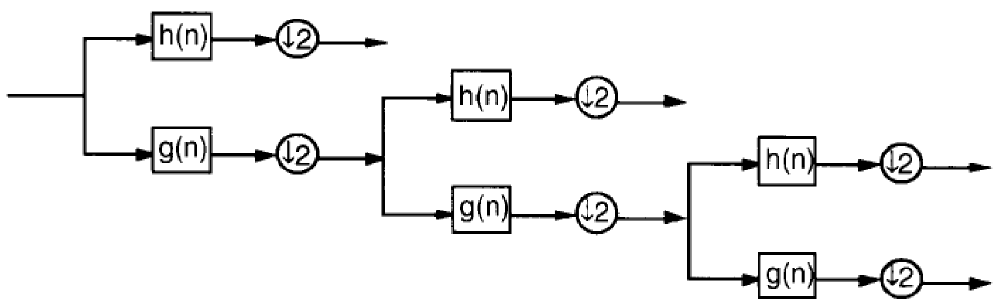

Figure 2: Illustration of a filter bank cascade. h(n) is an ideal lowpass filter, while g(n) is an ideal highpass filter.

2 Embedded Zerotrees of Wavelet Coefficients

Shapiro was one of the first authors to apply wavelet techniques to image compression [4]. In this paper, the 2-D wavelet transform was applied to the image by first subdividing the image into four equal subbands, determined using separable application of vertical and horizontal filters. The result is critically subsampled, such that each coefficient then corresponds to a 2x2 area of the image. The LH_1, HL_1, and HH_1 subbands are the finest scale coefficients, while the LL_1 subband is subdivided further in the very same fashion (see Figure 3). The subdivision continues until LL_N is a single coefficient, where N represents the number of subdivisions required for that operation. The result of the transform is the concatenation of the subband coefficients at each scale. This approach obviously parallels the cascading filter bank example discussed in the previous section, suggesting a tree structure to the subband decompositions.

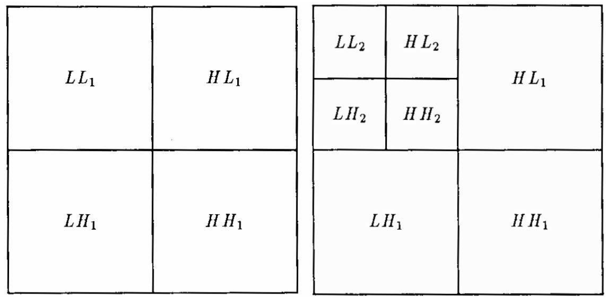

Figure 3: Left: The first stage of a 2-D wavelet transform. Right: after the second stage.

Given that an image has been transformed, the remaining problem is to decide the significance of each coefficient, and if a particular coefficient is insignificant, represent that information somehow. In this paper, the initial significance is decided by a simple thresholding operation. The insignificant coefficients are set to a particular value, zero.

The concept of a *zerotree* is based on the hypothesis that if a particular wavelet coefficient is insignificant, then all of the corresponding coefficients in finer resolution subbands will also be insignificant. This has some parallels to the JPEG method of using an end-of-block (EOB) code when the remainder of the DCT coefficients in a particular block are zero. However, this coder simply inserts a zerotree (ZTR) symbol when it detects an insignficant coefficient. The corresponding coefficients in finer subbands are then simply ignored. The method described in the paper uses multiple passes, with each pass using a decreasing value for the threshold. It also quantizes the significant coefficients in order to represent them in fewer bits. However, successive approximation is used to allow the significant coefficients to be quantized more precisely on successive passes. In this way the encoding can continue until a desired number of bits has been used to represent the image. The more bits used, the better the quality of the image.

The embedded zerotree wavelet (EZW) coder, as this was called, was determined to have significantly better PSNR performance than JPEG at low bit rates.

3 Set Partitioning in Hierarchical Trees

Probably the most successful variation of the ideas of Shapiro was Set Partitioning in Hierarchical Trees (SPIHT) [5]. SPIHT is based on the concept that wavelet coefficients with higher magnitudes should be transmitted first because they have more information content.

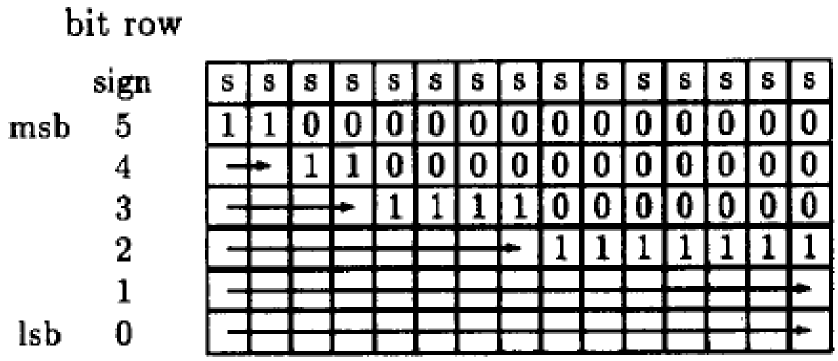

Figure 4: Left: The first stage of a 2-D wavelet transform. Right: after the second stage.

Figure 4 shows how a set of coefficients might be ordered and the information would be represented. The number of coefficients μ_n such that $2^n \leq |c| < 2^{n+1}$ is transmitted along with this information, where c is the coefficient in question. This means that only the bits covered by the arrows in Figure 4 need be transmitted because all other bits can be inferred from μ_n.

Although the coefficients might be ordered in terms of magnitude, in order to recover the image in the decoder, the coefficients will have to be reordered back to their original order. This ordering information is not transmitted explicitly, and instead is reconstructed based on the fact that the decoder duplicates the encoder's execution path. Each decision is denoted by

$$S_n(T) = \begin{cases} 1, & \max_{(i,j)\in T} |c_{i,j}| \geq 2^n, \\ 0, & otherwise, \end{cases} \tag{1}$$

This indicates the importance of the coordinates in T, where T is a continually updated set of coefficient coordinates used in the algorithm. At each encoding step, the S_n decision is output, and the decoding algorithm looks exactly the same except that S_n is input at each step of the way. This is what allows the execution path to be reconstructed and the ordering information inferred.

Instead of using zerotrees like EZW, SPIHT continually updates a set of internal buffers that contain coefficients that are in various stages of the algorithm. Because what is in these sets is updated internally with only the decision-making visible, the flow of information is more compact than in EZW. The performance is so much better that SPIHT before arithmetic coding is applied has similar performance to EZW, and only improves in performance once arithmetic coding is applied, with the additional cost of the encoding complexity.

SPIHT was significantly better in terms of SNR than any other published method at the time, and continues to be competitive with methods released more recently, which tend to be variants of SPIHT.

4 Tree-Structured Wavelets Applied to SAR

More recently, some authors have attempted to apply tree-structured wavelet techniques to synthetic aperture radar (SAR) imagery [6]. The proposed method is similar to the SPIHT encoding method described above, but it has a number of features to specialize it for SAR images. These include the notion of speckle reduction along with a special method of texture preservation. Figure 5 describes the block diagram of this method.

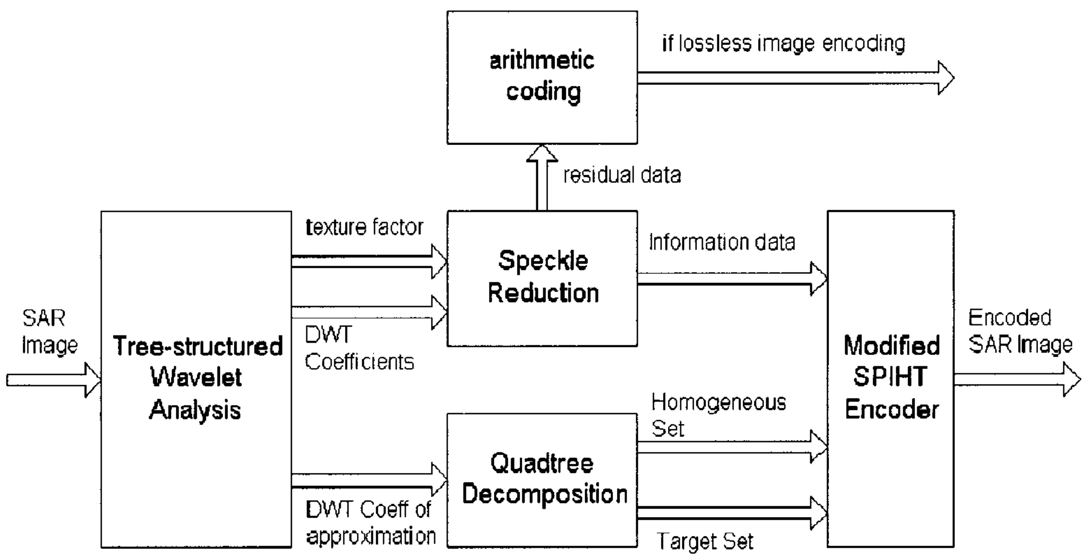

Figure 5: Block diagram of the tree-structured SAR data compression system

An important difference between SPIHT and the method used in this paper is that in the subband decomposition, not only the LL_n subband is decomposed at any given level. Because in SAR, preserving medium to high-frequency texture information is of particular concern, every subband is a candidate for further decomposition, with the energy of the subband being the metric for this decision. This is done by defining $F_l = KL/\sigma_l$, where F_l is a constant betwen 0 and 1, σ_l is the standard deviation (corresponding to the energy level) of a subimage, and K is a factor determined by the type of texture expected. The condition of Equation 1 in SPIHT is the changed to be $\max |c_{i,j}| \geq F_l 2^n, (i,j) \in T$. This allows texture information to have the highest bit priority.

Speckle reduction in this method is accomplished by performing a soft thresholding operation on the wavelet coefficients. This involves subtracting a value t_l from each coefficient, where $t_l = F_l \sigma_h \sqrt{2 \log n_l}$, and n_l is the number of pixels in each frequency band, and F_l is as define above.

The methods employed by this modified SPIHT were found to have good performance for the SAR images that were tested in the paper. The de-noising was found to have a positive effect on the SNR. The success of this method demonstrates that while SPIHT is a very efficient method for encoding the information in images, it is not the ultimate image coder by any means. There exists potential for improvement by customizing an image coder so that desirable information in a particular image type is preserved, while undesirable information is thrown away.

5 Conclusion

Issues involved in the compression of synthetic aperture radar signals have already been documented [7], [8]. These issues include high data rates coupled with relative scarcity of storage. The same principles that motivate compression in the radar case motivate it for sonar as well, although there may be some differences in the characteristics of the data.

While the performance of wavelet-based techniques for radar data have been reviewed, it is unknown if these techniques will apply equally as well to the sonar data in our posession. While both of these data sets are corrupted by speckle noise, which the algorithm in the last paper attempts to solve, it is unclear how important texture information is in these sonar images. In the sonar data, edge information is probably more important. In this case, I plan to adapt tree-based wavelet techniques to emphasize the preservation of edge information in sonar images. This can probably be implemented by somehow giving preference to certain types of high-frequency wavelet coefficients while still keeping speckle-reduction techniques in place. It is hoped that similar gains in performance as between the

SAR-modified SPIHT algorithm and the original SPIHT can be had by making these modifications. The potential for improvement over SPIHT for specialized images by specializing the coding method was demonstrated by this paper.

References

[1] A. Friedman, S. Mitchell, T. Kooij, and K. Scarbrough, "Circular synthetic aperture sonar design," in *Proc. Oceans*, vol. 2, June 2005, pp. 1038–1045.

[2] F. Sakarya, D. Wei, and S. Emek, "An evaluatin of SAR image compression techniques," in *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing*, vol. 4, April 1997, pp. 2833–2836.

[3] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 14–38, Oct. 1991.

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

[5] A. Said and W. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Syst. for Video Technol.*, vol. 6, pp. 243–250, Jun 1996.

[6] Z. Zeng and I. Cumming, "SAR image data compression using a tree-structured wavelet transform," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, pp. 546–552, Mar. 2001.

[7] J. W. Owens, M. W. Marcellin, B. R. Hunt, and M. Kleine, "Compression of synthetic aperture radar phase history data using trellis coded quantization techniques," in *Proc. IEEE International Conference on Image Processing*, vol. 1, Oct. 1997, pp. 592–595.

[8] G. Arslan, M. Valliappan, and B. Evans, "Quality assessment of compression techniques for synthetic aperture radar images," *Proc. IEEE International Conference on Image Processing*, vol. 3, pp. 857–861, Oct. 1999.