The Compression of Synthetic Aperture Sonar Images

Thomas Higdon

May 9, 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 image compression techniques, applies them to a given sonar data set, and assesses the performance of these methods. The Set Partitioning in Hierarchical Trees (SPIHT) method is found to be a good candidate, with reasonable complexity and excellent performance.

1 Introduction

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 paper will be to introduce the concepts necessary for the understanding of the image compression techniques that are applicable to sonar. Another objective will be to present applications of the wavelet transform to image compression, including academic journal papers on Set Partitioning in Hierarchical Trees (SPIHT). The final objective will be to apply these techniques to the sonar imagery data set that we have and assess the results.

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]. It is expected that a similar trend will be seen for sonar images. To this end, the use of the wavelet packet transform in image compression will be investigated, as it is hypothesized that high-frequency edges present in the data will be better represented by this generalization of the wavelet transform.

2 Background

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). The DCT 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, but 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. [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.

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.

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 1. 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.

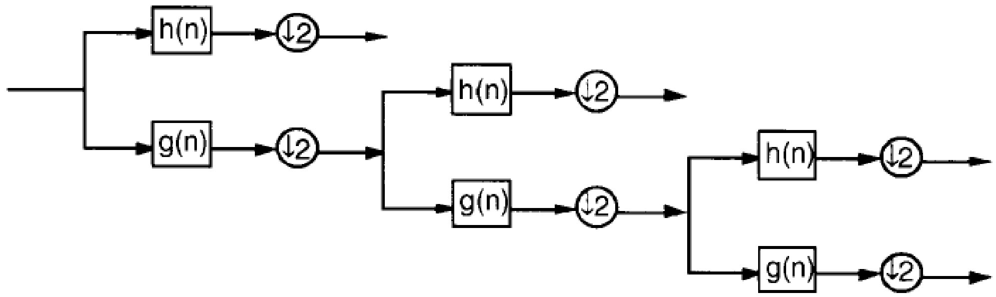

Figure 1: Illustration of a filter bank cascade. h(n) is an ideal lowpass filter, while g(n) is an ideal highpass filter.

2.2 Set Partitioning in Hierarchical Trees

Probably the most successful variation of the idea of wavelet zerotrees [4] 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.

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 the embedded zerotree wavelet technique (EZW) [4], 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 PSNR 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.

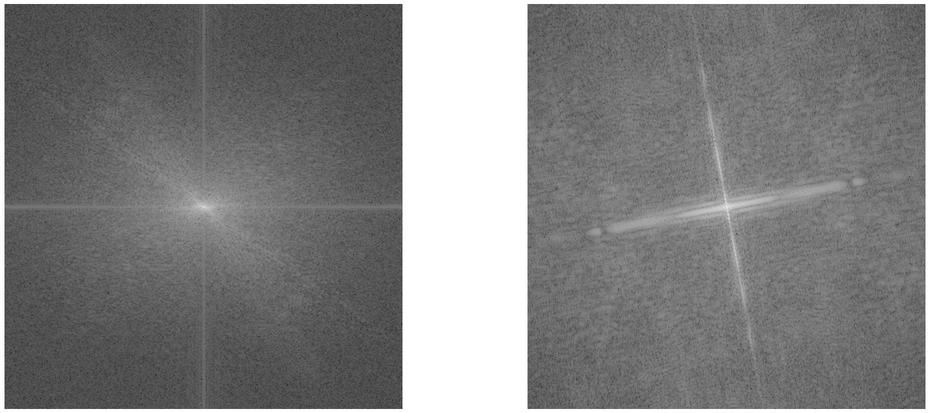

Figure 2: Left: Magnitude Spectrum of “Lena”. Right: Magnitude spectrum of sonar “cylinder” image.

3 Wavelet Techniques Applied to Sonar Images

It was thought that because the wavelet techniques like SPIHT have applied so well to optical images that they would also have some utility for sonar data images. The sonar image that was studied has some different frequency characteristics than a typical optical image. Figure 2 compares the magnitude of the frequency spectrum of the sonar image with the well-known optical “Lena” image. This figure demonstrates that there is more energy (and therefore more information) in the higher frequencies of the sonar image than in the those of the optical image. This difference in the frequency characteristics suggest that in the case of the wavelet transform, it might be useful to further subdivide the higher frequency subbands in search of greater information content for more efficient compression.

The wavelet packet transform (WPT) is a method for accomplishing this goal. The WPT differs from the wavelet transform method of packet decomposition in that the WPT continually subdivides each frequency subband, rather than just the lower ones. While the LL frequency subband contains most of the image information the vast majority of the time, it’s thought that further subdivision of the other frequency subbands can yield more efficient storage of information if certain entropy characteristics are met.

A full wavelet packet decomposition to the lowest level reveals redundant information and is computationally intensive, especially for larger images. A fast method [6] for selecting a useful wavelet packet basis has been found. This method uses a given rate-distortion criterion to select subbands to be divided based on the number of bits needed to approximate and image with a given error.

In [7], the authors describe a technique based on the previously described wavelet packet

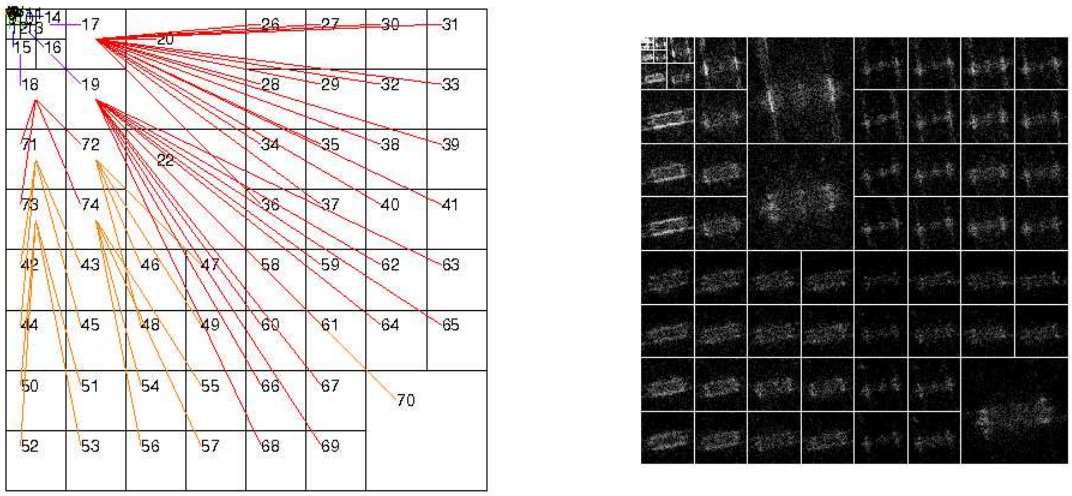

Figure 3: Left: Illustration of a subband hierarchy in wavelet packet-based SPIHT. Right: The wavelet packet coefficients for the sonar image.

expansion, and add to it their method for coding based on SPIHT. In traditional SPIHT, the subband hierarchy is easily described by the algorithm in that a "child" (a set of subband pixels at a finer level) always has a single "parent" (a subband pixel at a coarser frequency level). To use SPIHT techniques with wavelet packets, it is necessary to define what the hierarchy will look like for all of the possible cases, which [7] was able to do.

As can be seen from Figure 3, for the sonar "cylinder" image, the wavelet packet transformation has decomposed the higher frequency wavelet subbands in a large number of cases. The fact that these subdivisions have taken place indicates the possibility for greater compression beyond what is possible in the case of the ordinary wavelet transform. Wavelet packet transform-based techniques typically perform well on images that have rapidly oscillating components. Fingerprint images are an oft-used exampled. The "cylinder" image has this characteristic in that two of its edges are double-edges, a consequence of the manner in which the acoustic waves reflect from the physical target. It is hoped that the WPT-based technique could represent this phenonmenon better than than the wavelet-based technique.

For the WPT-based SPIHT technique, code [8] based on the results of [7] was used. This code implements the WPT-based SPIHT techniques described in the paper. For the JPEG2000 code, a JAVA-based implementation that interfaced with MATLAB was used.

Image quality assessment for this application is tricky. Some techniques, like Peak Signal to Noise Ratio (PSNR), have no basis in actual perceptual quality, only measure some quantitative difference between the reference and compressed image. Others, like the Universal Quality Index (UQI) and Weighted Signal to Noise Ratio (WSNR), have been developed using models of the

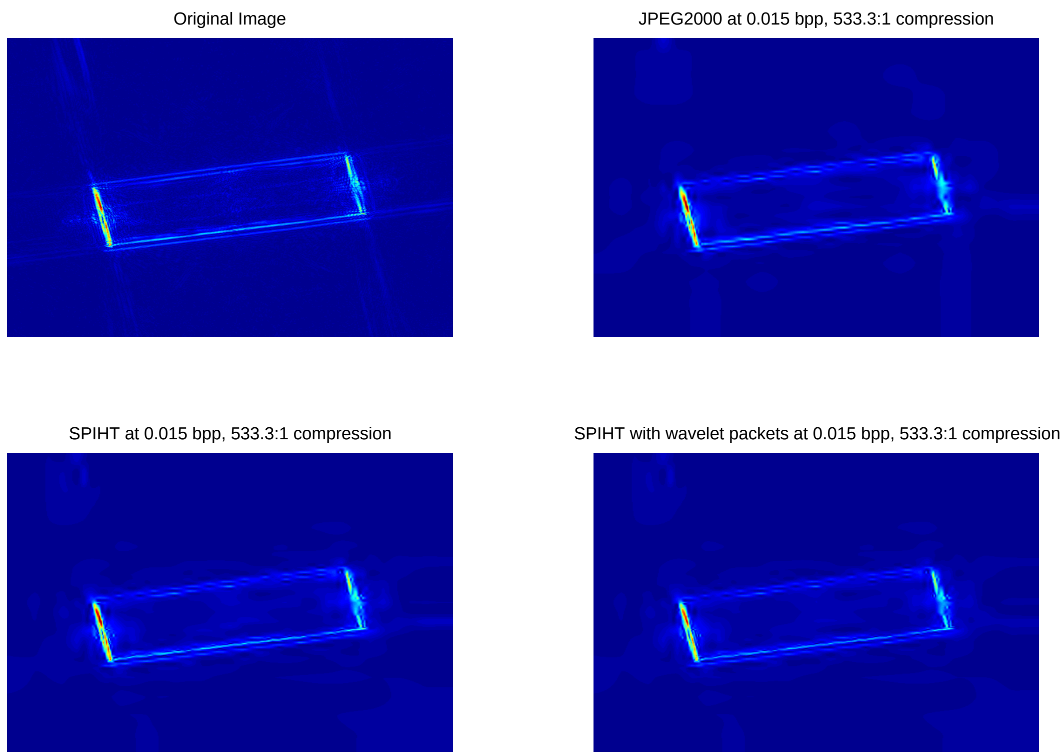

Figure 4: Demonstration of image compressors on the sonar image at 0.015 bits per pixel.

human visual system, and are specifically suited for human-consumed images. Sonar images are typically consumed by machine algorithms. In order to develop a better image quality metric for this type of image, more detail would have to be known about the eventual consumer of the image. Unfortunately, for this paper, this information was not known, and PSNR along with human visual system-based metrics were used for lack of better candidates.

4 Results

It can be seen in Figure 4 that the subjective quality of both the traditional SPIHT and wavelet-based SPIHT outdistance the JPEG2000 compressed image. However, little difference can be seen between the two SPIHT-based techniques. In all three compressed images, the edges can easily be distinguished, however there has been significant blurring, especially in the JPEG2000 image. None of the techniques were very accurate in reproducing the rightmost edge, where in all cases, there was a significant blob-like artifact and lack of definition to the edge.

Figure 5 illustrates some figures of merit for each of the image compression techniques tested

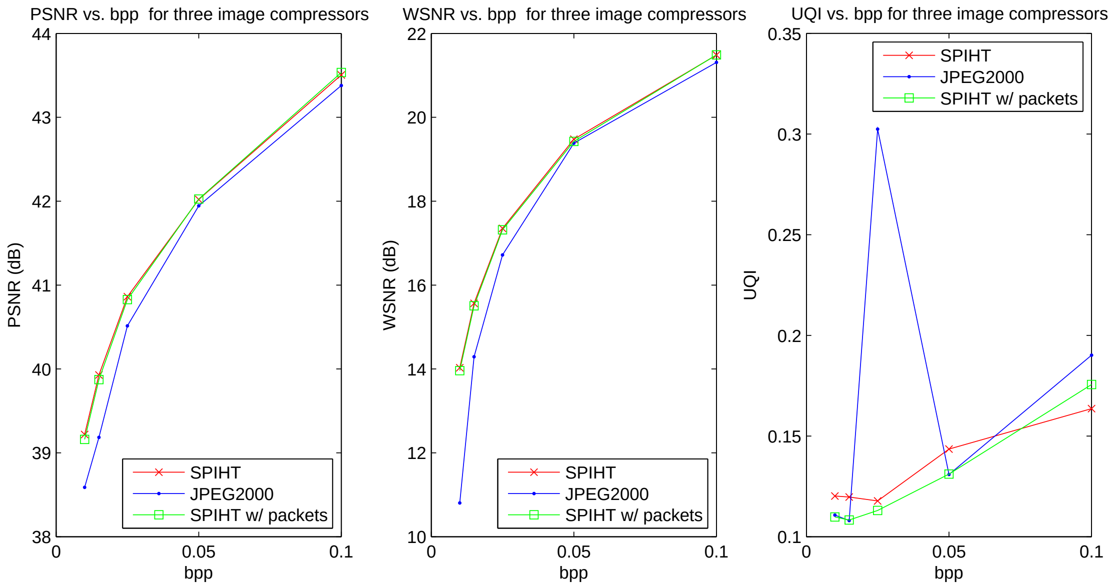

Figure 5: Left: Peak Signal to Noise Ratio (PSNR) of the tested image compression schemes. Center: Weighted Signal to Noise Ratio (WSNR) Right: Universal Quality Index (UQI).

at various bit rates. PSNR and WSNR show similar results to what one would have guessed given a look at the images, with SPIHT and SPIHT with wavelet packets tied, and JPEG2000 slightly behind. SPIHT with packets slightly edges out SPIHT at the higher bit rates for these metrics. Interestingly, the UQI results show some anomalous readings for JPEG2000. For the SPIHT-based techniques, there is a more pronounced difference between the two. The trend is the same, though, that the wavlet packet-based technique has slightly higher quality at the higher bit rate.

5 Conclusion

Issues involved in the compression of synthetic aperture radar signals have already been documented [9], [10]. 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.

The question posed by this project was whether a wavelet-packet based technique is capable of producing better compression for sonar images. The answer is that while wavelet packets seem to provide a slight edge in the image metrics used at some bit rates, the increased complexity required by the WPT-based techniques are probably not worth it.

An interesting direction for further research would be to customize the SPIHT technique to some of the newer image quality metrics other than PSNR or MSE. Traditional SPIHT assumes that the wavelet coefficients with the highest magnitude convey the most information because they reduce the mean-squared error of the final image the most. It would be interesting to see if using some model of UQI or WSNR in the wavelet domain would allow SPIHT to optimize its choice of wavelet coefficients to encode first to output the best image based on that metric.

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 evaluation 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] F. Meyer, A. Averbuch, and J. omberg, "Fast adaptive wavelet packet image compression," *IEEE Trans. on Image Processing*, vol. 9, 2000.

[7] N. Sprljan, S. Grgic, and M. Grgic, "Modified SPIHT algorithm for wavelet packet image coding," *Real-Time Imaging*, vol. 11, no. 5-6, pp. 378–388, 2005.

[8] N. Sprljan, "MATLAB zerotree toolbox (v1.21) [`http://www.sprljan.com/nikola/matlab`]," 2008.

[9] 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.

[10] 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.