EECS551 Deterministic Signal Processing Project Paper

Wavelet-Based Image Compression Considerations

Eric Durant
Electrical Engineering and Computer Science Department
University of Michigan
Ann Arbor, MI, USA
<edurant@umich.edu>

Monday 14 December 1998

Table of Contents

EECS551 – Deterministic Signal Processing Project Paper

Wavelet-Based Image Processing Considerations

Abstract

Wavelet-based image compression topics, both established and recent, are investigated, drawing on research papers, books and a technical article. Software for research and commercial applications is briefly reviewed.

Introduction

Focus is put on recent research in wavelet-based image compression. The rate-distortion tradeoff is a central thread. Lagrangian optimization and dynamic programming are discussed as efficient optimizations of this tradeoff.

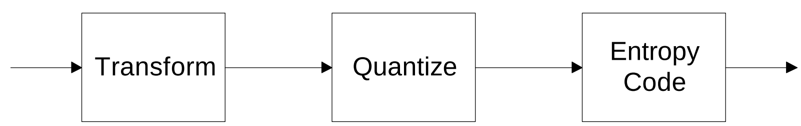
Other topics include encoder structure, model complexity tradeoff, motivation for transform coding, extension of the wavelet transform to the 2-D case, reasons wavelets provide a good representation of typical images, choice of wavelet bases, progressive coding, choice of a wavelet packet decomposition tree, relation to traditional coding theory and quantization.

The applied topics of standards and encoder / decoder tradeoffs are discussed. Finally, two software tools (one for research and the other for commercial use) implementing wavelet-based image compression are investigated.

Theory

General Encoder Structure

Figure 1[1] shows the general structure of an image compression system. The transform block moves image data to a domain more appropriate to compression than the spatial domain. In the quantize block significant transform coefficients are selected and represented with a specific number of bits. The entropy code block uses information theoretic principles to output a high entropy / low redundancy representation of the image.



• Figure 1: General Encoder Structure

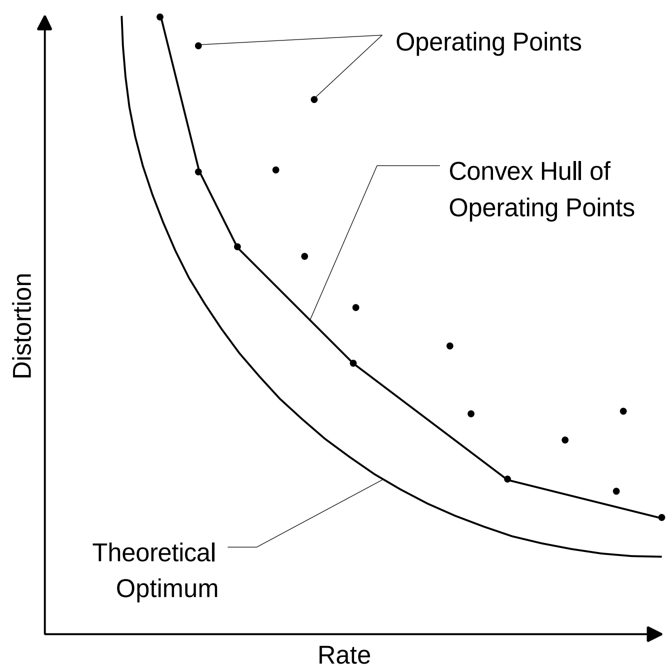
This paper focuses primarily on the transform, but significant issues regarding quantization are also discussed. Entropy coding is not investigated.

Rate-Distortion Tradeoff

Compressing an image with any parametric technique (e.g. JPEG/DCT, wavelets or wavelet packets) is an optimization problem balancing rate against distortion (this is the rate-distortion, or R-D tradeoff). Rate is generally taken as bits per pixel (bpp). Distortion, often taken as mean squared error (MSE), is more elusive. Although MSE does not necessarily correlate with perceptual quality, it is often the best objective measure of distortion. However, especially given a proper model (see “Model Complexity Tradeoff” below), MSE and subjective distortion do correlate well [Ort98, 24].

[1] Adapted from [Bur98]

Referring to Figure 2[2], a number of operating points (points representing a specific R-D tradeoff) are generated by various choices of compression system parameters. None of these can perform better than some theoretical optimum bound for a particular image. The operating points form a convex hull, on which all points may or may not be achievable in practice (see “Rate-Distortion Optimization” below).



• Figure 2: Operating Points and R-D Tradeoff

Model Complexity Tradeoff

Before an image is coded, an image model must be chosen. Early image coding techniques often considered frequency subbands independent, and thus were unable to exploit a common type of redundancy found in many image classes [Ort98, 26]. This is similar in concept to the Fourier transform, which moves a signal from the time (or spatial) domain to the frequency domain. Often, mixed

[2] adapted from [Ort98, 28]

time-frequency or time-scale (wavelet) transforms better capture the underlying structure of image data, forming a basis for more efficient compression in the R-D sense.

The tradeoff for a more robust model is increased computational complexity (consider the time difference between performing an FFT and a wavelet or time-frequency transform in MatLab).

There are many model choices and development techniques. For example, complex (and hence more capable) models may be realized by combining simpler models [Ort98, 28]. A major class of models comes from transform coding, of which wavelet decompositions are instances.

Transform Coding Motivation

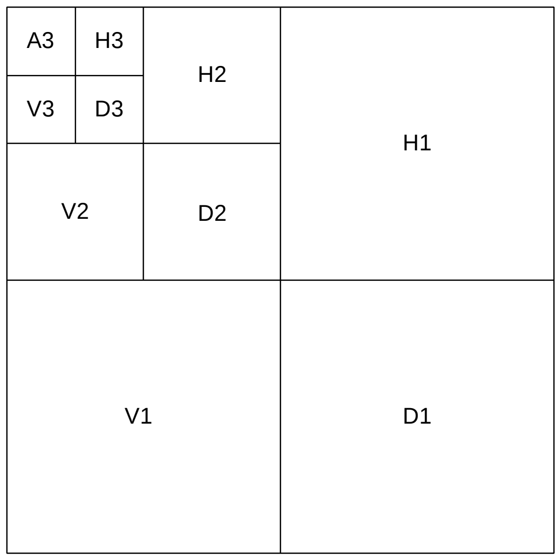
Transform coding is representing the signal to be compressed in an alternate domain (e.g. space-scale instead of Cartesian space for an image). R-D optimization is then performed in the transform domain. By design, this transform should have two properties for the class of signals to be compressed. First, the transform should concentrate the majority of the signal energy into a small number of coefficients (thus having minimal distortion for low rates). Second, the transform should result in decorrelated coefficients. This allows a simple scalar quantizer to have performance near that of a more complicated vector quantizer [Ort98, 29].

Generalization of Wavelets to 2-D Data

Figure 3 shows a 3-level instance of a 2-D wavelet decomposition proposed by Mallat, with higher decomposition indices representing lower scale features [Ant92]. The 2-D scaling and wavelet functions are given by:

$$\phi(x, y) = \phi(x)\, \phi(y) \qquad \psi^{H}(x, y) = \phi(x)\, \psi(y)$$

$$\psi^{V}(x, y) = \psi(x)\, \phi(y) \qquad \psi^{D}(x, y) = \psi(x)\, \psi(y)$$



• Figure 3: 2-D Wavelet Coefficient Organization

Wavelet / Image Affinity

Wavelet decompositions are efficient for a large number of common image properties such as gentle background gradients (scale localization) and edges (spatial localization) [Ort98, 30].

For many images, energy is concentrated in the lower scale subimages. Also, there is normally redundancy between subimages (both within a scale and between scales) which can be exploited by algorithms such as those using Sharpiro’s zerotrees [Mal98].

More flexibility and more efficient decompositions can often be found by generalizing to a wavelet packet decomposition optimized for a particular image or image segment [Ort98, 30]. This is the major topic of [Ram93] (see “Choice of Wavelet Packet Decomposition Basis” below).

Choice of Wavelet Decomposition Basis

Several properties are desired in a wavelet basis for images. Some of the most important are [Ant92]:

- smooth reconstruction filter (since smooth areas dominate most images)
- short filters (for efficient computation)
- linear phase filters (which lead to simpler cascade structures)

Since the only symmetric exact reconstruction (orthonormal) filters come from the Haar basis, a biorthogonal system with linear phase is preferred.

Expanding on smoothness, Antonini et. al. [Ant92] found that both vanishing moments of the decomposition wavelet and regularity of the reconstruction wavelet are important in improving both subjective and objective compression measures. In many cases, increasing the reconstruction regularity, even at great expense in decomposition vanishing moments, improves results.

Sufficient vanishing moments in the decomposition wavelets and regularity of the reconstruction wavelets used with vector quantization eliminate the blocking effects of using vector quantization in the spatial domain. Additionally, such wavelets do not have undesirable ringing effects [Ant92].

Progressive Stream Structure

Often, images are encoded for the primary purpose of efficient, on demand transmission over a data network. When the time to receive and decode an image stream is significant (more than a fraction of a second to more than a few seconds depending on the source consulted and the user's expectations) a progressive bitstream structure may be preferred. This allows the image to be visually rendered as it is received, normally with low scale or low frequency features first, greatly mitigating the perceived channel delay. Normally there is a cost in aggregate rate for arranging the data this way, but it often small (or insignificant) compared to the perceptual performance gains. This is precisely the rationale behind progressive JPEG encoding.

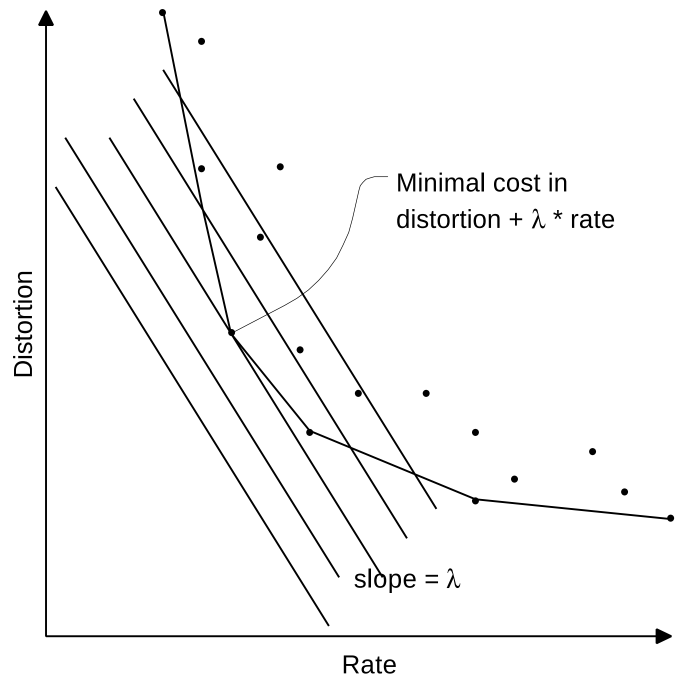
Maldonado-Bascón et al. [Mal98] have developed and implemented a progressive bitstream method for wavelet-based image compression. The authors point out that the drawback of their method is that it does not take advantage of the correlation between the subimages (e.g. zerotrees). The method begins by sending mean values for the subimages and then sends the largest deviations (in the image domain MSE sense), working its way down to the smaller deviations. For most images, this is roughly equivalent to a low to high scale transmission order. For two standard test images (Lincoln and Lena), at rates from 0.1 to 0.5 bpp, their algorithm performed about 1.0 to 1.5 dB worse in PSNR than Said and Pearlman's algorithm, which they recognize as the best published wavelet algorithm. This can be taken as an approximation of the distortion cost for structuring a stream progressively.

Rate-Distortion Optimization

The R-D optimization problem is choosing the optimal (in some sense) point in the operating space for a particular coder (see Figure 2). One technique for this is Lagrangian optimization, which is fast but cannot find the optimal point in some situations. Another R-D optimization technique is dynamic programming, which is much more computationally intensive, but more flexible than Lagrangian optimization.

Lagrangian Optimization

Lagrangian optimization reduces R-D optimization to specifying a single parameter, λ, the Lagrange multiplier, which captures the importance of low rate compared to the importance of low distortion for a particular application. Lagrangian optimization is illustrated in Figure 4. It can be envisioned with a series of parallel lines of slope λ in the R-D plane, each representing equal cost. The optimal solution (minimal cost) occurs where a line is tangent to the convex hull of operating points. For a given λ, the optimum quantization for each tile in a time-scale plane may be found this way. [Ort98, 38]



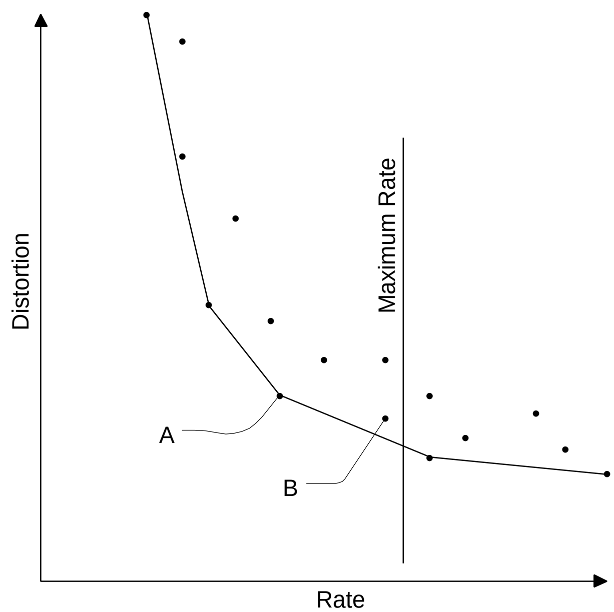
• Figure 4: Lagrangian Optimization

Dynamic Programming

Dynamic programming transverses the operating points in a tree structure. It is exhaustive and can therefore find better solutions than Lagrangian optimization under certain constraints. During transversal, if to solutions have the same rate but different distortion, the path with the higher distortion is pruned [Ort98, 39].[3] Requiring the same rate implies that there are a small number of possible rates (e.g. allocation of 20 bits to various subbands). This is consistent with using dynamic programming for sparse spaces and Lagrangian optimization for dense spaces, as is suggested by the example in the following paragraph.

[3] Since we expect the low order coefficients to be the most significant, we would not expect allocating the same number of bits to higher order of coefficients to change the relative distortion of the two paths. This is consistent with the nesting of subspaces in the wavelet transform.

Consider Figure 5[4], representing the allocation of a fixed aggregate number of bits (maximum allowable rate) to various time-scale channels of an image decomposition in which minimum distortion is desired. Lagrangian optimization can only find points on the convex hull and therefore will select operating point A, while dynamic programming will find B, the best tradeoff for the given operating points and rate limitation. In many applications, operating points are densely packed, essentially eliminating the benefits of dynamic programming [Ort98, 41].



• Figure 5: Dynamic Programming

Choice of Wavelet Packet Decomposition Basis

Certain images, such as those with high scale stationary components, are not well localized and decorrelated by the standard logarithmic wavelet decomposition tree. In such cases, adaptive methods, such as selecting the optimal wavelet packet decomposition tree, may yield better results. The most efficient wavelet packet tree in the Lagrangian sense (which is a function of the

[4] Adapted from [Ort98, 42]

Lagrangian parameter λ) can be found by starting with a complete wavelet packet decomposition tree and pruning nodes from the bottom. A pair of nodes is pruned if its parent node (representing a superspace of each child) can be encoded to achieve a better R-D tradeoff [Ram93].

Mismatch with Traditional Theory

Image data has much less stringent fidelity requirements than ASCII text and executable code in which one mistransmitted bit can potentially corrupt the entire message. [Ort98, 45] Thus, with proper encoder/decoder design, performance gains can be achieved by trading reliability for throughput (e.g. TCP vs. UDP). Also, since all bits in an image stream are not equally important (the tree structure is more important than the low order coefficients which are generally more important than the high order coefficients) a layered approach [Ort98, 46] in which metadata (e.g. tree structure) is sent on a more reliable transport than coefficient data can also increase performance.

Choosing a Quantization Scheme

Antonini et. al. [Ant92] investigated vector quantization. It can be more efficient than scalar quantization[5], and many of its common problems (e.g. not being based on psychovisual characteristics and having high computational cost) are mitigated by the wavelet transform.

Przelaskowski [Prz98] found that uniform scalar quantizers, which are typically near optimal for high bit rate applications, are inefficient for low bit rate applications. He investigated several quantization schemes and found that using

[5] The paper used *intra*-subimage vector codebooks. Given that correlation *between* subimages (both within and between levels) is common, it would be interesting to investigate whether a vector quantizer could exploit this correlation.

locally optimized quantizers for each subimage improved performance by about 1.0 dB PSNR over uniform threshold quantization for various 8-bit medical and standard test images at rates of 0.3 to 0.7 bpp.

Applications

Standards-Based Coding

To encourage commercially viable implementations of a coding method, a “bitstream syntax” is chosen so that encoders and decoders can be designed and optimized separately and be guaranteed to interoperate. This is the basis for the success of such standards as JPEG and MPEG [Ort98, 31].

Encoder / Decoder Tradeoff

There are several reasons to design bitstreams that prefer flexible (and hence more complex) encoders and simple decoders. A flexible encoder can take advantage of advances in theory without requiring replacement of existing decoders. Also, properly designed, a flexible encoder will have better performance on signals not well matched to the underlying model when compared to a simpler encoder [Ort98, 31].

Also, increasing the cost at the encoder often yields superior returns in decreased cost and simpler decoders, much as the pilot of an FM radio signal is wasteful of energy, but allows receivers to be built with great reliability and low cost. Ortega [Ort98, 34] summarizes this tradeoff well:

> Complex algorithms can be justified in scenarios where encoding is performed just once but decoding is done many times.

Results and Discussion

Two software programs were used to investigate wavelet coding of images. MatLab's Wavelet toolbox provides building blocks for theoretical investigation, while Summus' 4U2C provides a commercial application for image compression.

MatLab's Wavelet Toolbox

MatLab's Wavelet Toolbox allows viewing the results of wavelet and wavelet packet compression on images with various wavelets at user-selectable scales and can generate transform data. However, it does not provide a quantitative rate measure such as file size or bpp for the wavelet transform. It would be an interesting extension of the existing routines to write MatLab code to generate R-D graphs for a variety of wavelets, scales and thresholding schemes.

Summus' 4U2C

Summus' 4U2C wavelet-based image compression software allows interactive image compression by letting the user set compression ratio, quality factor (similar to standard JPEG encoders) or file size. Since it is a consumer product, details such as the bases and decomposition trees being used are not available.

Subjectively, the algorithm provided much better results than JPEG, especially at medium (about 1:40 with a highly detailed image, 24-bit image) to low bitrates (below 1:200). The encoding cost seems significantly greater than JPEG, though, with compression taking one to two seconds on hardware which can perform JPEG compression in a small fraction of a second. The relicing of low bitrate JPEG was replaced by a more acceptable loss of fine texture (e.g. bricks on a wall), but at a significantly lower bitrate than JPEG. The wavelet compression did an excellent job of maintaining overall color and tight outlines even at low bitrates (lower than 1:200 on a highly detailed image). At even lower bitrates (lower than 1:500, allocating only several hundred bytes to estimate a detailed image), the general forms were discernable and the overall colors of the image were well preserved.

Consistent with the low bitrate subjective results, using the demo images and browser plugins from the Summus Website (http://www.summus.com/), the progressive display feature provides much clearer results at lower scan size than progressive JPEG.

Conclusion

Wavelet-base image compression theory has rapidly progressed in the last several years. With the basic theory in place, software tools have become available which aid in further research and consumer applications of the technology. For example, the JPEG 2000 draft standard[6] incorporates wavelet compression technology. On the research side, MatLab's Wavelet Toolbox provides many building blocks for designing and testing wavelet image compression systems.

[6] Details and current revisions are only available to standards committee members, so a more thorough of the standard could not be provided in this paper.

References

- [Ant92] Antonini, M., Barlaud, M., Mathieu, P. and Daubchies, I., "Image Coding Using Wavelet Transform", IEEE Transactions on Image Processing, Vol. 1, No. 2, Apr. 1992, pp. 205-220.
- [Bur98] Burrus, C.S., Gopinath, R.A. and Guo, H., "Introduction to Wavelets and Wavelet Transforms: A Primer", Prentice-Hall, Upper Saddle River, NJ, USA, 1998.
- [Mal98] Maldonado-Bascón, S., López-Ferreras, F. and Amo-López, P., "Unhierarchical Coding of Image Wavelet Coefficients Without Sending the Sign", Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Oct. 1998, Pittsburgh, pp. 465-468.
- [Ort98] Ortega, A. and Ramchandran, K., "Rate-Distortion Methods for Image and Video Compression", IEEE Signal Processing Magazine, Vol. 15, No. 6, Nov. 1998, pp. 23-50.
- [Pra91] Pratt, W.K., "Digital Image Processing: Second Edition", John Wiley & Sons, New York, 1991.
- [Prz98] Przelaskowski, A., "Fitting Quantization Scheme to Multiresolution Detail Preserving Compression Algorithm", Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Oct. 1998, Pittsburgh, pp. 485-488.

- [Ram93] Ramchandran, K. and Vetterli, M., “Best Wavelet Packet Bases in a Rate-Distortion Sense”, IEEE Transactions on Image Processing, Vol. 2, No. 2, Apr. 1993, pp. 160-175.