JOINT OPTIMAL BIT ALLOCATION AND BEST-BASIS SELECTION FOR WAVELET PACKET TREES

Jill R. Goldschneider   Eve A. Riskin

Department of Electrical Engineering, Box 352500
University of Washington, Seattle, Washington 98195-2500
jrgold@ee.washington.edu riskin@ee.washington.edu

ABSTRACT

In this paper, an algorithm for wavelet packet trees that can systematically identify *all* bit allocations/best-basis selections on the lower convex hull of the rate-distortion curve is presented. The algorithm is applied to tree-structured vector quantizers used to code image subbands that result from the wavelet packet decomposition. This method is compared to optimal bit allocation for the discrete wavelet transform.

1. WAVELETS

Traditional Fourier analysis decomposes a signal into a sum of orthogonal trigonometric functions, creating a frequency representation of the data. However, Fourier analysis eliminates all temporal or spatial information. When compressing signals such as images, it is often desirable to capture local frequency characteristics, especially since a great portion of signal energy in the frequency domain is consumed by discontinuities. A *time-frequency* or space-frequency representation of a signal can be used to capture local frequency behavior of signals and images.

Wavelets [1–3] are an important tool for time-frequency analysis in which simple, orthonormal bases of $L^2(\mathbb{R}^d)$ are built with good localization properties in both space and frequency providing a powerful framework for image coding. Wavelets are used to transform an image into localized orthogonal components containing different band-pass frequency information. Localization is especially important for handling image features such as abrupt changes due to boundaries or edges.

To create orthonormal wavelet packets, a signal is decomposed into smooth (**S**) and detail (**D**) subbands corresponding to different orthogonal frequency bands. To create more packets, the subbands are further decomposed. An isotropic wavelet packet tree (WPT) decomposition [3] with orthogonal subbands is shown in Figure 1. Each signal space is the direct sum of its subspaces. Hence the original signal can be represented by the leaves of any subtree of the WPT. For example, the common discrete wavelet transform (DWT) is the subtree comprised of subbands **D**, **DS**, **DSS**, and **SSS**. Fast and efficient implementation of the wavelet packet transform can be done in $\mathbf{O}(N \log N)$ which is equivalent to the time required by the fast Fourier transform. The use of a predetermined subtree of the WPT, such as the DWT, may reduce computation time.

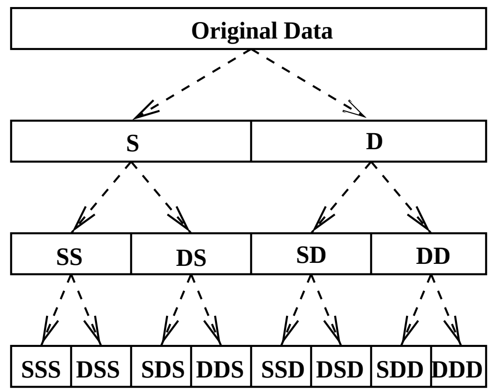

Figure 1: The wavelet packet decomposition for a 1-d signal decomposed to 3 levels. **S** and **D** refer to smooth and detail subbands.

The main advantage of wavelet packets is better signal representation by choosing the subbands that are better adapted to the frequency characteristics of the signal. The search for the "best" non-redundant representation of the data by any subtree of the WPT is called *best-basis selection* [4]. Best-basis selection begins by evaluating each subband with a desired metric (e. g. rate or distortion). Then, using this information, a post-ordered search of the WPT is done while a best-basis decision for each branch is made. The post-ordered traversal enables exploitation of the recursive property that each signal space is the direct sum of its subspaces. An example of this is given in Figure 2.

Two important problems associated with image compression are the quantization of the wavelet subbands and the allocation of bits to each subband. If using wavelet packets, best-basis selection must also be considered. An algorithm for joint bit allocation/best-basis selection which

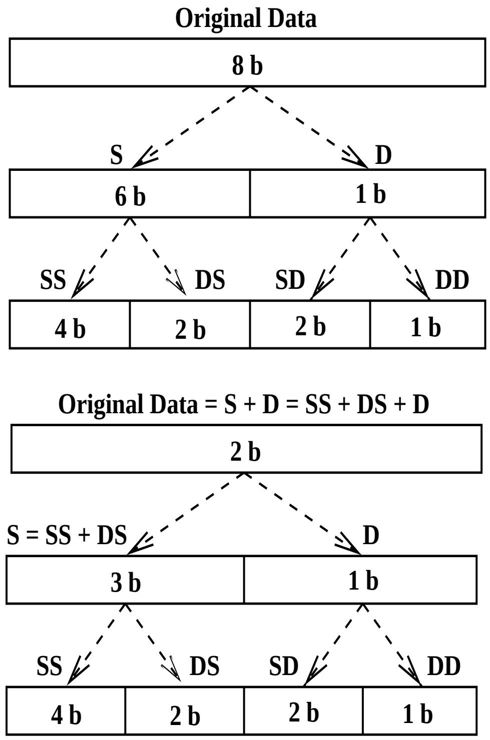

Figure 2: Best-basis selection for a 1-D 2-level WPT based on rate. (a) The WPT with initial rates. (b) Using a post-ordered search, the best-bases for subbands **SS**, **DS**, **SD**, and **DD** are the subbands themselves. For subband **S**, it is better to use subspaces **SS** and **DS** with average rate 3 b, than the subband itself which requires 6 b. For subband **D**, it is better to use subband **D** with average rate 1 b, than subspaces **SD** and **DD** which requires 1.5 b. Finally, for the original data, subspaces **SS**, **DS**, and **D** with average rate 2 b should be used rather than the original signal requiring 8 b.

heuristically finds a WPT bit allocation/best-basis selection on the lower convex hull of the rate-distortion curve close to the design rate was presented in [5]. However, due to the heuristic search, it is possible that it is not the solution closest to the desired rate. In this paper, a systematic approach to joint bit allocation/best-basis selection which will find all vertices on the lower convex hull of the rate-distortion curve is presented. By having the optimal rate-distortion characteristics for all bit rates, a precise description of the encoder alone can be used in the design of more complex encoding schemes.

2. VECTOR QUANTIZATION

Vector quantization (VQ) is a lossy compression technique that has been used extensively in speech and image compression [6]. VQ exploits the memory or correlation existing between neighboring signal samples by quantizing them together. Tree-structured VQ [6] (TSVQ) is a low-complexity alternative to full search VQ where the codebook is structured as a binary (or M-ary) tree and the codewords are leaves of the tree.

A well known approach to TSVQ design is to create a high-rate TSVQ, and then prune it based on rate-distortion trade-offs to yield a pruned TSVQ (PTSVQ) with the desired rate [7]. One common pruning algorithm is the generalized Breiman, Friedman, Olshen, and Stone (GBFOS) algorithm [8]. GBFOS PTSVQ's typically outperform both fixed rate TSVQ and full search VQ over most rates of interest [7]. GBFOS produces PTSVQ's that are optimal in the sense that they lie on vertices of the lower convex hull of the rate-distortion curve. In addition, GBFOS can systematically locate *all* codebooks on the lower convex hull, providing a complete description of the optimal rate-distortion characteristics of the TSVQ by rate, distortion, and slope of the vertices.

3. BIT ALLOCATION

Bit allocation is the process of assigning a given number of bits to the wavelet subbands to minimize the overall distortion of a coder. A good bit allocation method usually results in much better performance by devoting more bits to regions of the signal that are active or difficult to code and fewer bits to less active regions.

An orthonormal, non-redundant wavelet transform has the property that the average rate and MSE distortion of a signal quantized in the wavelet domain is the average rate and MSE distortion of the quantized coefficients. This is convenient for performing bit allocation to all subbands simultaneously. A subtree S of an orthonormal WPT with leaves $\tilde{S}$ offers a non-redundant representation with average rate and distortion

$$R(S) = \sum_{n \in \tilde{S}} R(n) \quad \text{and} \quad D(S) = \sum_{n \in \tilde{S}} D(n),$$

where $R(n)$ and $D(n)$ are the normalized rate and distortion of the encoded subband n.

To optimize bit allocation so that for a given rate the lowest distortion quantizer is used, the Lagrangian minimization technique is used to formulate the quantization cost as the average distortion plus a penalty for the rate of the subtree:

$$J(S) = \sum_{n \in \tilde{S}} D(n) + \lambda \sum_{n \in \tilde{S}} R(n).$$

The cost is minimized when $\lambda = -\frac{\partial D(n)}{\partial R(n)}$, regardless of the subtree. Therefore, optimal bit allocation is achieved when the quantizers for each subband lie on the same slope on the lower convex hull of the rate-distortion curve.

4. OPTIMAL BIT ALLOCATION AND BEST-BASIS SELECTION

For optimal bit allocation, the problem is to minimize the cost $J(S)$ for a given subtree S. For best-basis selection, the problem is to find the subtree S^{opt} which minimizes the cost over all other subtrees:

$$S^{\mathrm{opt}} = \arg\min_{S \subset T} \left[J(S) \right].$$

If λ is fixed, then $J(S)$ is the desired measure used to find the best-basis. Thus, by fixing λ, the best basis is identified by minimizing the cost, and then the rate and distortion are evaluated. This is the foundation for [5] which heuristically selects different values of λ until a rate close to the desired rate is found.

To systematically find all bit allocation/best-basis solutions on the lower convex hull, it is necessary to know all slopes on the lower convex hull ranging from 0 to ∞. Then the best bases, rates and distortions of the WPT at all vertex points can be evaluated.

Again, by using WPT quantizers which lie on the lower convex hull of their local rate-distortion curves, and which intersect the same slope, the best-basis selection is recursively computed by a post-ordered traversal of the WPT. This reduces the joint bit allocation/best-basis search to that of finding the lower convex hull of each WPT branch, that is, the lower convex hull of two convex hulls must be identified. An example is shown in Figure 3 where the solid line represents the lower convex hull of the given subband, the dotted line represents the lower convex hull of the best-bases of the node's children, and the gray line represents the lower convex hull of the optimal bit allocation/best-basis selection between the given node and all of its descendents.

Assuming that the rate-distortion curves for the subband PTSVQ's are found via GBFOS (i. e. the curves are convex, and the rate, distortion, and slopes of the vertices are known) then the slopes for the lower convex hull of the WPT can be easily found. If there is no change in best basis, then the slope of the WPT's lower convex hull is the same as the slope of the current best-basis. If there is a change in bases, then the slope of the WPT is the slope between the two vertex points being compared. The change in bases can be detected by computing the slope between the two current vertex points. If this slope is greater than the largest slope to the right of the two current vertex points, and less than the smaller slope to the left of the two current vertex points then there is a change in bases.

The change can also be detected by examining the cost function $J(S)$. Each vertex v can be intersected by a slope in the interval $[\lambda_l(v), \lambda_r(v)]$ where λ_l and λ_r refer to the slopes to the left and right of the vertex. Two vertices v_1, v_2 can be intersected by any slope in the interval

$$[min(\lambda_l(v_1), \lambda_l(v_2)), max(\lambda_r(v_1), \lambda_r(v_2))].$$

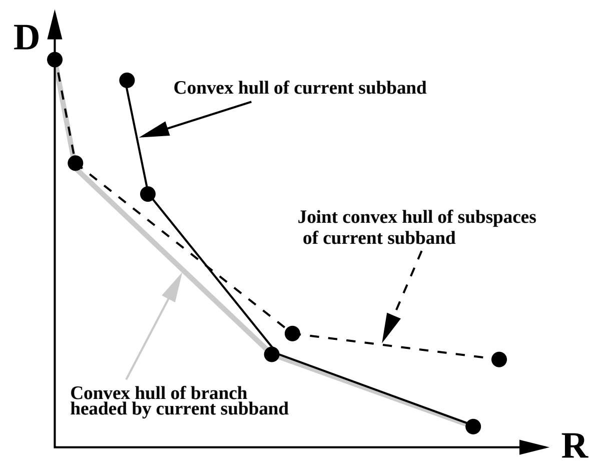

Figure 3: Optimal bit allocation/best-basis selection

If the cost function $J(S)$, evaluated at the two endpoints of the interval, shows that both vertices are on the lower convex hull, then a change of bases will occur.

Details of the algorithm are given in [9]. Regardless of how the change is detected, after each post-ordered traversal, the current best basis of each branch is described by its rate, distortion, and the minimum slope leading to the next adjacent vertex on the lower convex hull of the rate-distortion curve.

5. RESULTS

For these experiments, WPT/PTSVQ optimal bit allocation was done using USC database images crowd, couple, man, woman1, and woman2 as training data. The PSNR of the training data as a function of WPT depth (Figure 4) improves with increased WPT depth. This is expected as more levels of decomposition permits more adaptability in signal representation and compression. However, improvements decrease with increasing WPT depth, and it may not be worthwhile to use a WPT with more than 3 levels.

Bit allocation as a function of vector dimension is examined in Figure 5, where it is seen that it is better to use larger vector dimensions. Note that some curves start increasing rapidly at higher bit rates. This is a reflection of overtraining of the PTSVQ's due to the finite amount of training data. All of the curves exhibit this behavior although it is not shown in these graphs.

A comparison of DWT/PTSVQ and WPT/PTSVQ optimal bit allocation for three decomposition levels as a function of vector dimension is shown in Figure 6. As expected, WPT/PTSVQ outperforms DWT/PTSVQ for all vector dimensions since the DWT is a subset of the WPT.

6. CONCLUSIONS

In this paper, the problem of joint bit allocation and best-basis selection for wavelet packets was addressed by an algorithm that identifies all solutions on the lower convex hull of the rate-distortion curve. This algorithm was applied to PTSVQ's used to code image wavelet packet subbands, and compared to optimal bit allocation for the DWT. It was found that WPT/PTSVQ outperformed DWT/PTSVQ, as expected since the DWT is a subset of the WPT, and the WPT approach is inherently adaptive. It was also seen that results improved as the vector dimension and decomposition level increased.

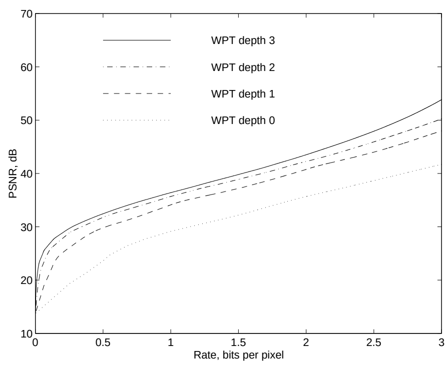

Figure 4: WPT/PTSVQ bit allocation as a function of WPT depth. The vector dimension is 4 for all PTSVQ's.

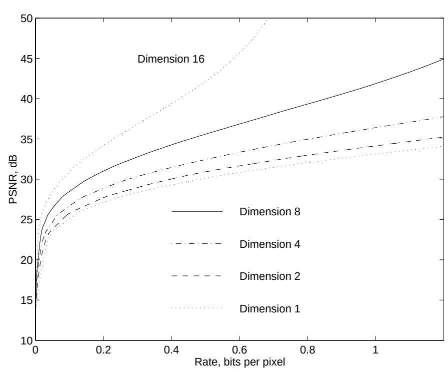

Figure 5: WPT/PTSVQ bit allocation as a function of vector dimension. All PTSVQ's in a WPT have the same vector dimension.

7. ACKNOWLEDGMENTS

This work was supported by a NASA Global Change Fellowship, an NSF Young Investigator Award, and U.S. Army Research Office Grant DAAHOU-96-1-0255.

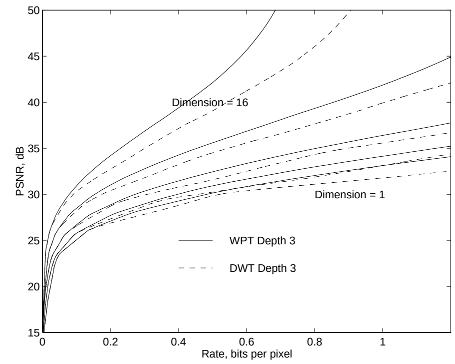

Figure 6: WPT/PTSVQ versus DWT/PTSVQ bit allocation for 3 levels of decomposition.

8. REFERENCES

[1] I. Daubechies, *Ten lectures on wavelets*. Philadelphia, PA: Society for industrial and applied mathematics, 1992.

[2] J. Kovačević and M. Vetterli, *Wavelets and Subband Coding*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.

[3] M. V. Wickerhauser, *Adapted Wavelet Analysis – from theory to software*. A. K. Peters, Ltd, 1994.

[4] R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 713–718, 1992.

[5] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 160–175, 1993.

[6] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer Academic Publishers, 1992.

[7] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," *IEEE Transactions on Information Theory*, vol. 35, pp. 299–315, Mar. 1989.

[8] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. The Wadsworth Statistics/Probability Series, Belmont, California: Wadsworth, 1984.

[9] J. R. Goldschneider, *Lossy Compression of Earth Science Data*. PhD thesis, University of Washington, Seatlle, WA, June 1997.