Appears in *Proceedings of the Data Compression Conference* (DCC), 2001, IEEE Computer Society Press.

The Coding-Optimal Transform

Cynthia Archer and Todd K. Leen
Department of Computer Science and Engineering
Oregon Graduate Institute of Science & Technology
20000 N.W. Walker Rd, Beaverton, OR 97006-1000
Email: archer, tleen@cse.ogi.edu

Abstract

We propose a new transform coding algorithm that integrates all optimization steps into a coherent and consistent framework. Each iteration of the algorithm is designed to minimize coding distortion as a function of both the transform and quantizer designs. Our algorithm is a constrained version of the LBG algorithm for vector quantizer design. The reproduction vectors are constrained to lie at the vertices of a rectangular grid.

A significant result of our approach is a new transform basis specifically designed to minimize mean-squared quantization distortion for both fixed-rate and entropy-constrained coding. For Gaussian distributed data, this transform reduces to the Karhunen-Loeve transform (KLT). However, in general the coding optimal transform (COT) differs from the KLT enough to provide up to 1 dB improvement in compressed signal-to-noise ratio (SNR) on images. We describe a practical algorithm that finds the COT for a given signal. In addition, we present image compression results demonstrating the SNR improvement achieved with our algorithm relative to KLT based transform coding.

1 Introduction

Transform coding is a low-complexity alternative to vector quantization and is widely used for image and video compression. A transform coder compresses multi-dimensional data by first transforming the data vectors to new coordinates and then coding the transform coefficient values independently with scalar quantizers. A key goal of the transform coder is to minimize compression distortion while keeping the compressed signal representation below some target size. While quantizers are typically designed to minimize compression distortion [1, 2], this is not the case for the transform. The coordinate transform has been fixed apriori, as in the discrete cosine transform (DCT) used in the JPEG compression standard [3]. The transform has also been adapted to the signal statistics using the Karhunen-Loeve transform (KLT) as in recently published transform coding work [4, 5]. These transforms are not designed to minimize compression distortion, nor are they designed (selected) in concert with quantizer development. For instance, the design goal of the KLT is to concentrate signal energy into a few components.

In this paper, we present a new algorithm for transform coder design that concurrently optimizes both transform and quantizers. Our algorithm is a constrained version of the Linde-Buzo-Gray (LBG) algorithm for vector quantizer design [6]. A significant result of our approach is a new transform basis designed to minimize mean squared compression distortion. In this paper, we derive the conditions this coding-optimal transform (COT) must satisfy to minimize distortion. In addition, we describe a simple algorithm for determining the transform. We conclude by presenting results from image compression experiments that compare the compression performance of COT-based transform coders with KLT-based transform coders.

2 Optimal Transform Coding

A transform coder converts a signal to new coordinates and then codes the transform coefficients independently of one another with scalar quantizers. One can think of a transform coder as a vector quantizer with the M reproduction vectors constrained to lie at the vertices of a rectangular grid. The grid is defined by orthogonal axes, s_J, $J = 1 \ldots n$ and n sets of scalar reproduction values, one for each dimension. There are M_J possible reproduction values on the s_J axis, thus the total number of grid vertices is $M = \prod_J M_J$. Encoding a n-dimensional data vector with a vector quantizer requires $\mathcal{O}(Mn)$ add/multiply operations for the distance calculations and $\mathcal{O}(M)$ compare operations. A transform coder requires $\mathcal{O}(n^2)$ add/multiply operations for the transform and naively $\mathcal{O}(\sum_J M_J)$ compare operations. However, efficient binary search techniques can be used to encode the scalar transform coefficients reducing the number of compare operations to $\mathcal{O}(\log_2 M)$.

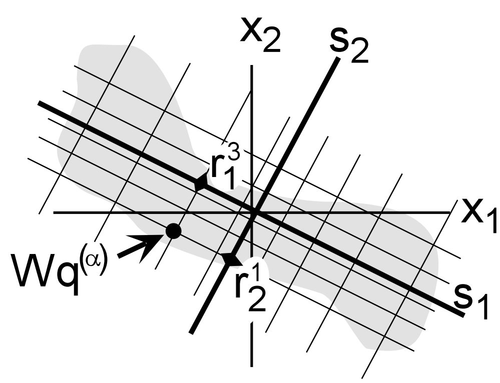


Figure 1: Orientation of quantizer grid in signal space. The quantizer reproduction vectors $q^{(\alpha)}$, $\alpha = 1 \ldots M$, lie at the vertices of a rectangular grid. The grid is oriented to the signal vectors x (indicated by the gray area) with orthogonal transform, W.

The compression/restoration process replaces each signal vector with one of a small set of reproduction vectors. The encoder assigns the transform coefficients of a data vector to codewords. The decoder replaces each codeword with the associated reproduction value. Figure (1) illustrates the structure of a two-dimensional transform coder. The r values indicate the scalar reproduction values; $r_J^{(i)}$ is the i^{th} value along the s_J axis. The coordinates of the reproduction *vectors*, $q^{(\alpha)}$, $\alpha = 1 \ldots M$ are combinations of the scalar reproduction values $[r_1^{(i)}, r_2^{(j)}, \ldots, r_n^{(k)}]^T$, $i = 1 \ldots M_1$, $j = 1 \ldots M_2$, etc. A reproduction vector $q^{(\alpha)}$ represents all the data vectors in region $R^{(\alpha)}$ of the data space. We will refer to the regions defined by the assignment of signal values to reproduction values as the *partition*.

The $n \times n$ orthogonal transform, W, defines the orientation of the quantizer grid in the data space. In the data coordinate basis, the reproduction vectors are given by $Wq^{(\alpha)}$. Conversely, in the transform basis, the data vectors are $s = W^T x$.

To develop an optimal transform coder for a particular signal, one finds values for the transform coder parameters that minimize compression-induced distortion. We quantify distortion with the mean-squared difference between the original and compressed signal vectors,

$$D = \sum_{\alpha=1}^{M} \int_{R^{(\alpha)}} \|x - Wq^{(\alpha)}\|^2 \, p(x) \, \mathrm{d}^n x \tag{1}$$

where x are signal vectors with density $p(x)$. The transform coder parameters are the orthogonal transform W as well as the bit allocation, reproduction values, and data partition that define the quantizer grid. Because each parameter is dependent on other parameters, we use an iterative optimization approach. We fix all but one parameter value and then minimize (1) with respect to that free parameter. Below we first discuss transform optimization and then quantizer optimization.

2.1 Transform Optimization

To optimize the transform, we find the orientation of the quantizer grid which minimizes distortion (1). The transform W is constrained to be orthogonal, that is $W^T W = \mathbf{I}$. The cost function to be minimized is thus

$$C = \sum_{\alpha=1}^{M} \int_{R^{(\alpha)}} \|x - \sum_{J=1}^{n} W_J q_J^{(\alpha)}\|^2 \, p(x) \, \mathrm{d}^n x + \sum_{K=1}^{n} \sum_{L=1}^{n} \gamma_{KL} (W_K^T W_L - \delta_{K,L}) \tag{2}$$

where W_J is the J^{th} column vector of W, q_K is the K^{th} coordinate of reproduction vector q, and γ_{KL} is a Lagrange multiplier. Minimizing C with respect to the transform matrix element W_{KJ} yields

$$\frac{\gamma_{JK} + \gamma_{KJ}}{2} + \sum_{\alpha} p^{(\alpha)} q_K^{(\alpha)} q_J^{(\alpha)} = \sum_{\alpha} q_J^{(\alpha)} \int_{R^{(\alpha)}} x^T \, p(x) \, \mathrm{d}^n x W_K = \sum_{\alpha} q_K^{(\alpha)} \int_{R^{(\alpha)}} x^T \, p(x) \, \mathrm{d}^n x W_J \tag{3}$$

where $p^{(\alpha)} = \int_{R^{(\alpha)}} p(x) \, \mathrm{d}^n x$. If we define the outer-product matrix Q

$$Q = \sum_{\alpha} q^{(\alpha)} \int_{R^{(\alpha)}} x^T \, p(x) \, \mathrm{d}^n x, \tag{4}$$

then (3) requires $QW = W^T Q^T$. This symmetry condition along with the orthogonality condition uniquely defines the coding optimal transform W.

By using the conditions for the coding optimal transform, we can determine this transform for two cases of interest, Gaussian data and high-resolution coding. Gersho and Gray [7] and Mallat [8] have shown, by using high-resolution distortion approximations, that the optimal coding transform for Gaussian data is the KLT. Using (3) it is possible to show that this is the case, *regardless of bit-rate*[1]. The product of Q and W is given by

$$QW = \sum_{\alpha=1}^{M} q^{(\alpha)} \int_{R^{(\alpha)}} s^T \, p_s(s) \, \mathrm{d}^n s \tag{5}$$

where $s = W^T x$ and W is orthonormal. We need two results to show QW is symmetric when W is the KLT. First we note that for Gaussian $p_x(x) = \mathcal{N}(0, \Sigma)$, W diagonalizes the covariance Σ, hence $p_s(s)$ is the product of marginals $\prod_J p_J(s_J)$. Second, the reproduction values which minimize mean-squared distortion are given by

$$q_K^{(\alpha)} = \frac{\int_{R_K^{(\alpha)}} s_k \, p_K(s_K) \, \mathrm{d}s_K}{\int_{R_K^{(\alpha)}} p_K(s_K) \, \mathrm{d}s_K} \tag{6}$$

where $R_K^{(\alpha)}$ is the projection of $R^{(\alpha)}$ onto the s_K axis. By substituting these two results into (5), it is straightforward to show that QW is symmetric, hence the KLT is the coding optimal transform when the data is Gaussian. Note that the partition (encoder) need not minimize mean squared error, so this result applies to entropy-constrained and uniform quantizers, as well as fixed-rate quantizers.

In the case of high-resolution coding, the reproduction values are so numerous and closely spaced that the data density in each region $R^{(\alpha)}$ is uniform, $p_x(x | x \in R^{(\alpha)}) =$ constant. If in addition, the reproduction values are given by minimum error quantizers (6), QW is symmetric for *any* orthogonal W. Consequently, in the high-resolution limit, distortion is independent of the orientation of the quantizer grid.

2.2 Quantizer Optimization

To determine the quantizer parameters, one minimizes distortion subject to a constraint on the number of coding bits. The number of coding bits can be constrained in two ways, limiting the average rate (entropy-constraint) or limiting the number of coding bits per vector (fixed-rate). For entropy-constrained compression, each scalar reproduction value, $r_J^{(i)}$, has an associated code length, $l_J^{(i)}$. The average rate $L = \sum_J \sum_i p_J^{(i)} l_J^{(i)}$, where $p_J^{(i)}$ is the probability that $s_J \in R_J^{(i)}$, is constrained to a target rate, B. For fixed-rate compression, each vector is coded with the same number of bits. The number of coding bits per vector, $L = \sum_J \log_2 M_J$, is constrained to be B bits. We discuss these two cases briefly below. For a recent comprehensive review of quantization see [10].

[1] For a alternate approach to this proof, see [9].

For *entropy-constrained* lossy compression, the average rate L should equal target rate B. In addition, for the code words to be uniquely decodable, the code lengths must satisfy the Kraft inequality, $\sum_i 2^{-l^{(i)}} \leq 1$. The cost function to be minimized, written in terms of the transform coefficients $s = W^T x$, is

$$\begin{aligned} C &= \sum_{J=1}^{n} \sum_{i=1}^{M_J} \int_{R_J^{(i)}} (s_J - r_J^{(i)})^2 \, p_J(s_J) \, \mathrm{d}s_J + \\ & \lambda \left(\sum_{J=1}^{n} \sum_{i=1}^{M_J} p_J^{(i)} \, l_J^{(i)} - B \right) + \sum_{J=1}^{n} \gamma_J \left(\sum_{i=1}^{M_J} 2^{-l_J^{(i)}} - 1 \right) \end{aligned}$$

where $R_J^{(i)}$ defines the range of coefficient values represented by $r_J^{(i)}$ and λ and γ_J, are Lagrange multipliers. By allowing the codes to have non-integer lengths, minimizing C with respect to the code lengths yields $l_J^{(i)} = -\log_2 p_J^{(i)}$. Minimizing C with respect to the reproduction values places $r_J^{(i)}$ at the mean of the transform coefficients in $R_J^{(i)}$ (6). Finally, minimizing C with respect to the partition of the data space yields regions

$$R_J^{(j)} = \{ s_J \,|\, (s_J - r_J^{(j)})^2 - \lambda l_J^{(j)} < (s_J - r_J^{(k)})^2 - \lambda l_J^{(k)} \quad \forall\, k \neq j \} \tag{7}$$

where the Lagrange multiplier λ is selected to enforce the rate constraint.

In *fixed-rate* compression the code lengths for the reproduction values in each quantizer are the same, $l_J^{(i)} = \log_2 M_J$. The number of coding bits per block, $L = \sum_J \log_2 M_J$, is constrained to be less that some target rate B. The cost function to be minimized is

$$D = \sum_{J=1}^{n} \sum_{i=1}^{M_J} \int_{R_J^{(i)}} (s_J - r_J^{(i)})^2 \, p_J(s_J) \, \mathrm{d}s_J + \lambda \left(\sum_{J=1}^{n} \log_2 M_J - B \right) \tag{8}$$

where $R_J^{(i)}$ defines the range of coefficient values represented by $r_J^{(i)}$ and λ is a Lagrange multiplier. Allocating the coding bits B where they minimize distortion (8) determines the optimal values for M_J, $J = 1 \ldots d$ [11, 12]. Minimizing distortion with respect to the reproduction values places the $r_J^{(i)}$ at the mean of the transform coefficients in $R_J^{(i)}$ (6). Finally, minimizing (8) with respect to the partition yields regions

$$R_J^{(j)} = \{ s_J \,|\, (s_J - r_J^{(j)})^2 < (s_J - r_J^{(k)})^2 \quad \forall\, k \neq j \} \tag{9}$$

3 Implementation

The algorithm for optimal transform coder design is a constrained version of the Linde-Buzo-Gray (LBG) algorithm for vector quantizer design [6]. It alternates between improving the transform and improving the quantizers until the constrained distortion measure reaches a local minimum.

3.1 Coding Optimal Transform Algorithm

The COT algorithm finds the orientation of the current quantizer grid that minimizes compression distortion (1). We initialize the grid orientation to the KLT[2]. At each iteration, we calculate the QW matrix from the transform coefficients and the reproduction values. To minimize distortion, we must find the W that makes the QW matrix symmetric (3). We quantify how far the matrix is from symmetric with the sum squared difference between transposed matrix elements

$$A = \sum_{K=1}^{n-1} \sum_{J=K+1}^{n} (a_{KJ} - a_{JK})^2 . \tag{10}$$

where a_{KJ} is the K^{th} row and J^{th} column element of QW. We apply Givens rotations [13], $G(K, J, \theta)$, to minimize A. Multiplication by $G(K, J, \theta)$ applies a rotation of θ radians to the (K, J) coordinate plane. For a $n \times n$ matrix, there are $\frac{n^2-n}{2}$ such planes. Minimizing (10) with respect to rotation $G(K, J, \theta)$ yields a solution for θ that is quartic in $\tan\theta$. However, when the angle is small, so that $\tan^2\theta \ll 1$, the solution simplifies to

$$\tan\theta \approx \frac{(a_{KK} + a_{JJ})(a_{KJ} - a_{JK}) - \sum_{I \neq K,J}(a_{JI}a_{IK} - a_{JI}a_{IK})}{\sum_{I \neq K,J}(a_{JI}a_{IJ} + a_{KI}a_{IK}) + (a_{KK} + a_{JJ})^2 - (a_{KJ} - a_{JK})^2} \tag{11}$$

In image compression experiments, we consistently found that the rotation angles were small. We find the rotation angle (11) for each coordinate plane and apply these rotations to the current transform matrix. This process is *repeated* until $A/\|QW\|_F$, where $\|QW\|_F$ is the Frobenius norm, is less than a threshold ($A \approx 0$). The new W will orient the quantizer grid so that compression distortion is minimized.

3.2 Quantizer Algorithms

Quantizer optimization defines scalar quantizers that represent the data with minimal distortion given a constraint on the compressed bit-rate. To develop the quantizers, we first transform the signal vectors to the the current transform basis W. Again we consider both entropy-constrained and fixed-rate compression cases.

For *entropy-constrained* compression, each quantizer is trained to optimally represent the transform coefficients using an entropy-constrained quantizer algorithm [2]. We initialize with ten-bit uniform quantizers. If the entropy $H = \sum_J \sum_i p_J^{(i)} \log_2 p_J^{(i)}$ is too far from the target rate ($|H - B| > 0.1$ bit), we adjust the Lagrange multiplier, λ, which enforces the rate constraint. The change in λ is $(B - H)/\frac{\partial H}{\partial \lambda}$ where $\frac{\partial H}{\partial \lambda}$ is estimated from the previous two values of H and λ. For the k^{th} adjustment, $\frac{\partial H}{\partial \lambda}^{(k)} = (H^{(k-1)} - H^{(k-2)})/(\lambda^{(k-1)} - \lambda^{(k-2)})$ with $\lambda^{(0)} = 0$ and $\lambda^{(1)} = 1$. The quantizers

[2]The quality of the final solution is very sensitive to the bit allocation determined at the initial quantizer grid orientation. To insure a good starting bit allocation, we always initialized W to the KL transform.

are then retrained with the new λ and the entropy is re-evaluated. This process repeats until the rate constraint is satisfied.

For *fixed-rate* compression, we use Riskin's [12] bit allocation[3] to determine the quantizer sizes. We maintain one through ten bit quantizers for each coordinate. Each quantizer is trained to optimally represent the transform coefficients using the Lloyd algorithm [1]. One then calculates the distortion for each quantizer size and coordinate. Starting from zero bits in each coordinate, one allocates one, two, or more bits at a time to the coordinate where the additional allocation will reduce the distortion per coding bit the most. This method results in an allocation of coding bits that is at or close to the desired number of bits B and that is on the convex hull of possible rate-distortions.

4 Experimental Results

We illustrate the difference between the KLT and COT using two-dimensional data that is sampled from two intersecting Gaussian distributions[4]. Figure 2 contains a plot of this data overlaid with a one by two bit quantizer grid. The KLT aligns the grid along the dominant high-variance Gaussian, consequently data from the lower variance Gaussian is poorly represented. The COT rotates the quantizer grid so that the reproduction vectors better represent all the data. The compressed data signal-to-noise ratio (SNR) is 0.46 dB higher when the COT orients the quantizer.

We also exercised our transform coders on image data. In plots 3 and 4, we show SNR results for two classic test images, Barbara and Goldhill[5]. The plots in figure 3 are for entropy-constrained compression; entropy coding was not performed. The plots in figure 4 are for fixed-rate compression.

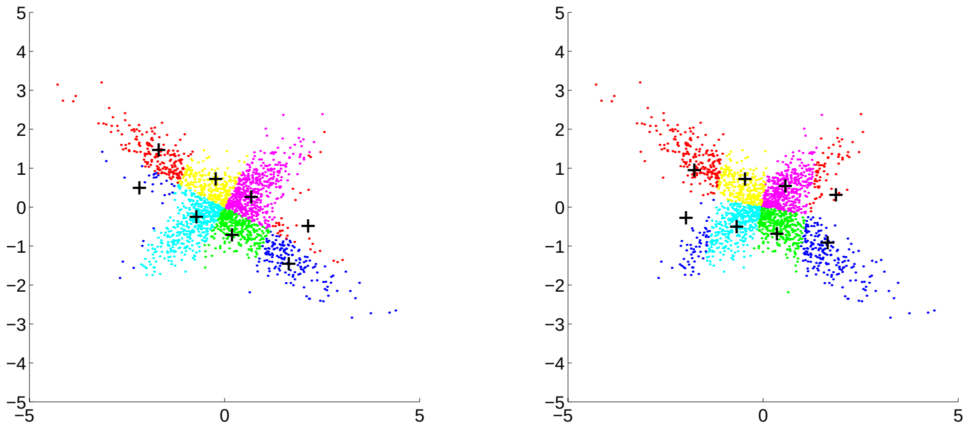
For *entropy-constrained* compression, our experiments show that using the COT instead of KLT increases SNR by 0.3 to 1.2 dB for entropies in the range of 0.25 to 1.25 bits per pixel (bpp). Of the images tested, Barbara showed the largest SNR improvement when the COT is used and Goldhill showed the smallest improvement. Other tested image types (e.g. frames from natural image video, magnetic resonance images) showed similar SNR improvements.

For *fixed-rate* compression and low bit rates, using the COT instead of KLT increases image SNR very little. The high bit-rate COT basis vectors are essentially the same

[3] In [12], Riskin defines two allocation methods. We use the method that does *not* require convexity of the rate-distortion function.

[4] 2000 data points were sampled from a mixture of two Gaussians, $\mathcal{N}(0, U_1^T \Sigma_1 U_1)$ with $U_1 = \begin{bmatrix} -.6 & .8 \\ .8 & .6 \end{bmatrix}$ and $\Sigma_1 = \begin{bmatrix} 4 & 0 \\ 0 & .16 \end{bmatrix}$ and $\mathcal{N}(0, U_2^T \Sigma_2 U_2)$ with $U_2 = \begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & .16 \end{bmatrix}$.

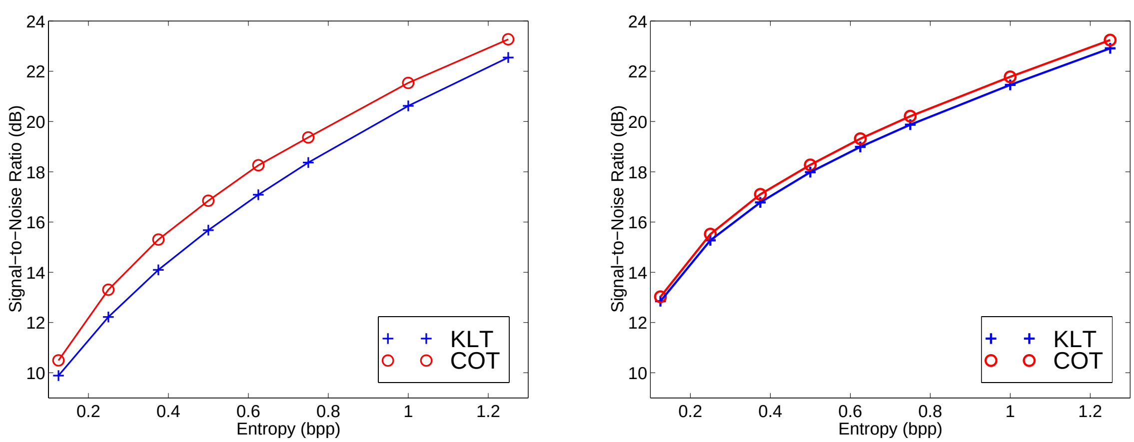
[5] Barbara is a photograph of a seated women wearing striped clothing. Goldhill is a photograph of a row of houses in a hillside village. These images are available from the University of Waterloo website, http://links.uwaterloo.ca/greyset2.base.html.



(a) KLT (b) COT

Figure 2: The quantizer on the left is oriented with the KLT, the one of the right with the COT. Data vectors are indicated with ·'s. and the reproduction vectors are indicated with +'s.

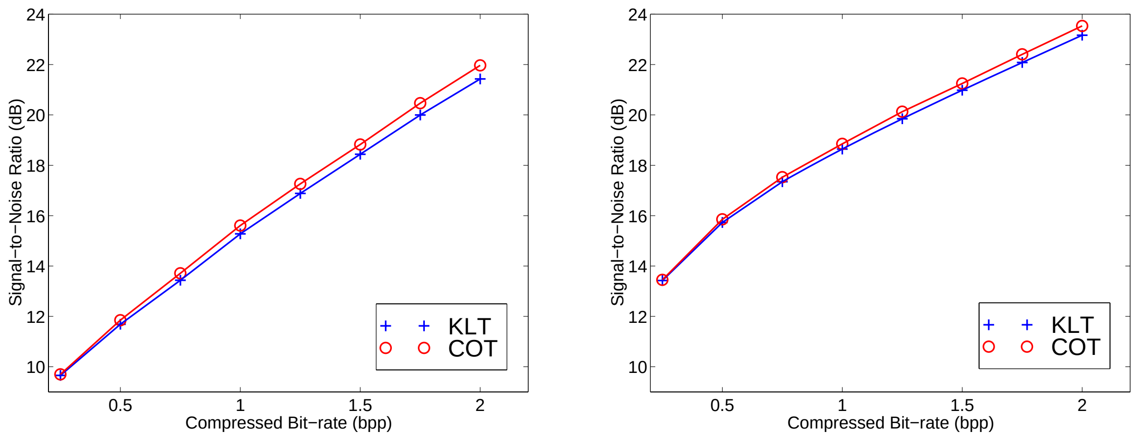
as the high-variance KLT basis vectors, so when only a few coordinates are coded there is little difference in SNR. However, the SNR improvement due to using the COT increases as more coordinates are coded, since for image data the mid-variance KLT basis vectors differ from the mid and low bit-rate COT vectors. At 1.0 bpp orienting the quantizer grid with the COT instead of KLT increases SNR by 0.2 to 0.35 dB.



(a) Barbara (b) Goldhill

Figure 3: Entropy-constrained compression: SNR versus entropy for Barbara and Goldhill test images.

COT-based transform coding is no worse than KLT-based coding in terms of storage overhead and encode/decode time. The storage overhead, which includes storing and transmitting the transform matrix and quantizer reproduction values, is the same for



(a) Barbara (b) Goldhill

Figure 4: Fixed-rate compression: SNR versus bit-rate for Barbara and Goldhill test images.

both methods. Encoding and decoding times are also the same. However, in our variable-rate compression experiments, the COT-based coders required typically 3 to 4 times longer to train than did the KLT-based coders. For the Barbara image, the KLT-based coders required 130 to 135 seconds to train on a Sun SPARC Ultra2 with approximately 95% of the training time due to developing the quantizers. The COT-based coders required 225 to 680 seconds, depending on the number of training iterations. Transform optimization accounted for 30% to 60% of the training time.

5 Discussion

Transform coders are often constructed by concatenating an ad hoc choice of transform with bit allocation and quantizer design. Instead, we treat transform coder design as an optimization problem and derive a locally optimal algorithm. This algorithm is a constrained version of the LBG algorithm for vector quantizer design, with reproduction vectors constrained to lie at the vertices of a rectangular grid. This algorithm (as well as an adaptive version) can be derived from suitably constrained, probabilistic data models [14].

The derivation leads to a new transform basis, the coding optimal transform (COT), which unlike the KLT, is specifically designed to minimize compression distortion. Variable-rate image compression experiments show that using our COT instead of the KLT increases SNR by 0.3 to 1.2 dB. We have shown that the COT reduces to the KLT for Gaussian sources.

Like the KLT, the COT is a data dependent transform. Consequently, it suffers from the same drawbacks as the KLT; the transform must be calculated from the input signal and stored with the compressed signal. Because of these drawbacks, we expect COT-based transforms coders to be most effective in an adaptive or universal transform coding framework. An adaptive transform coder consists of several different transform coders, each optimized to compress a different signal type. Each signal vector is compressed using the transform coder that represents it with the least distortion. Our transform coding algorithm could be combined with the *adaptive* framework developed by Effros, et al. [5] to form a constrained LBG algorithm for adaptive transform coding.

Acknowledgements

We thank the reviewers for helpful comments. This work was supported by NSF under grant ECS-9976452.

References

[1] S. Lloyd. Least square optimization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[2] N. Farvardin and J. Modestino. Optimum quantizer performance for a class of non-gaussian memoryless sources. *IEEE Transactions on Information Theory*, IT-30(3):485–497, 1984.

[3] G. Wallace. Overview of JPEG (ISO/CCITT) still image compression standard. *Communications of the ACM*, 4(4):30–40, 1991.

[4] R. Dony and S. Haykin. Optimally adaptive transform coding. *IEEE Transactions on Image Processing*, 4(10):1358–1370, 1995.

[5] M. Effros, P. Chou, and R. Gray. Weighted universal image compression. *IEEE Transactions on Image Processing*, 8(10):1317–1328, 1999.

[6] Y. Linde, A. Buzo, and R.M. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28(1):84–95, January 1980.

[7] A. Gersho and R. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic, 1992.

[8] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[9] V. Goyal, J. Zhuang, and M. Vetterli. Transform coding with backward adaptive updates. *IEEE Transactions on Information Theory*, 46(4):1623–1633, 2000.

[10] R. Gray and D. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, 1998.

[11] Y. Shoham and A. Gersho. Efficient bit allocation for an arbitrary set of quantizers. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(9):1445–1453, 1988.

[12] E. Riskin. Optimal bit allocation via the generalized BFOS algorithm. *IEEE Transactions on Information Theory*, 37(2):400–402, 1991.

[13] G. Golub and C. Van Loan. *Matrix Computations*. John Hopkins University Press, 1989.

[14] C. Archer and T.K. Leen. From mixtures of mixtures to adaptive transform coding. In Leen, Dietterich, and Tresp, editors, *Advances in Neural Information Processing Systems 13*. MIT Press, 2001 (to appear).