

High-Rate Transform Coding: How High is High, and Does it Matter?

Vivek K Goyal

Bell Labs, Lucent Technologies, Murray Hill, NJ 07974-0636

v.goyal@ieee.org, <http://cm.bell-labs.com/who/vivek/>

Abstract — Karhunen–Loève transforms (KLT’s) are the optimal orthogonal transforms for transform coding of Gaussian sources. This well-known fact is usually established with approximations from high-resolution quantization theory. How high does the rate have to be for these approximations to be accurate? The minimum rate allocated to any component should be at least about one bit. (The average rate per component may be much higher.) Does the rate actually have to be high for the KLT to be optimal? No, the KLT is optimal more generally. Two new, simple proofs of this fact are described. They rely on a scale invariance property, but not on high-resolution approximations or properties of optimal fixed-rate quantization.

Let $\{x^{(n)}\}_{n \in \mathbb{Z}^+}$ be a sequence of independent, identically distributed (i.i.d.), zero-mean Gaussian random vectors of dimension N with covariance matrix $R_x = E[xx^T]$. In transform coding, an orthogonal linear transform T is applied to each source vector to get a vector of *transform coefficients* $y = Tx$. The transform coefficients undergo fixed- or variable-rate scalar quantization, yielding $\hat{y} = Q(y)$; a reproduction vector is obtained by inverting the transform: $\hat{x} = T^{-1}\hat{y}$. The fidelity of reproduction is measured by the mean-squared error per component between the source vector and the reproduction: $D = N^{-1}E\|x - \hat{x}\|^2$.

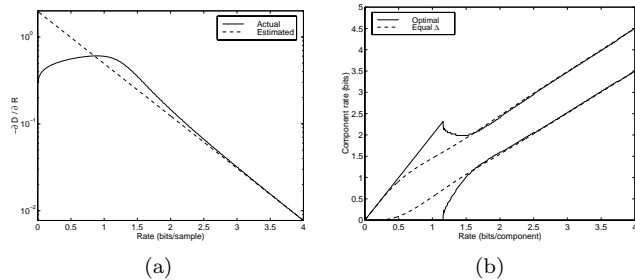
A transform that makes the transform coefficients uncorrelated is called a Karhunen–Loève transform (KLT). The optimality of the KLT was first shown by Huang and Schultheiss under assumptions of optimal fixed-rate quantization and a mild, commonsense condition on the bit allocation. (Earlier work by Kramer and Mathews did not involve quantization and was not in an operational rate–distortion framework.) Optimality of the KLT can also easily be established under the assumption that each component quantizer has distortion-rate performance described by

$$D_i = \frac{\pi e}{6} \sigma_i^2 2^{-2R_i}, \quad (1)$$

where σ_i^2 is the variance of y_i . This result relies on optimal, arbitrary-real bit allocation, which is unrealistic.

At high rates, (1) is a good approximation of the performance of entropy-coded uniform quantization (ECUQ). The original intention of this work was to determine how high the rate has to be for the KLT to be optimal or nearly optimal when using ECUQ. Actually, there is no limitation on the rate for the KLT to be optimal. Also, through numerical calculations, bit allocations based on (1) are close to optimal when each coefficient has a rate of at least one bit per sample.

Limits of high-resolution analysis Lagrangian bit allocation using (1) is easy because of the simple form of $\partial D_i / \partial R_i$. Where (1) is accurate, the optimal allocation of bits results in equal quantization step sizes for each transform coefficient.



The accuracy of the derivative of (1) is assessed in Fig. (a). With $\sigma_1^2 = 1$ and $\sigma_2^2 = 1/4$, optimal bit allocations are compared to those obtained with equal quantization step sizes in Fig. (b).

Optimality of the KLT The optimality of the KLT holds much more generally than previously published results indicated. The new result below does not rely on optimal fixed-rate quantization or high-resolution quantization theory.

Theorem [1] Assume that the distortion–rate performance of a scalar quantizer applied to a component with variance σ^2 is $D = \sigma^2 f(R)$. Then a KLT is an optimal transform, i.e., for any given maximum rate, it minimizes the distortion.

We may assume that $f(\cdot)$ is nonincreasing; if $R_1 > R_2$ but $f(R_1) > f(R_2)$, rate R_1 can be replaced in any purportedly optimal solution by rate R_2 . $f(\cdot)$ need not be convex.

Proof 1: Let T be any orthogonal transform. Suppose that R_i bits are allocated to transform coefficient y_i . Assume $\sigma_i^2 > \sigma_j^2$ implies $R_i \geq R_j$; otherwise, the distortion can be reduced by the permutation of T that swaps y_i and y_j .

If the (i, j) component of $R_y = TR_x T^T$ is nonzero for some $i \neq j$, the Jacobi rotation that zeroes this value does not increase the distortion. Repeating the process until convergence (the classical Jacobi algorithm for computing eigendecompositions) yields a KLT at least as good as T . ■

Proof 2 (Telatar): This proof is based on elementary properties of majorization [2]. The problem is to minimize the function $D = N^{-1} \sum_{i=1}^N \sigma_i^2 f(R_i)$ by manipulating the σ_i^2 's through the choice of T . Let $\sigma = (\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2) = \text{diag}(TR_x T^T)$. For a Hermitian matrix, the diagonal elements are majorized by the eigenvalues, so σ is majorized by a vector λ of eigenvalues of R_x . Now the majorization of σ by λ is equivalent to σ being in the convex hull of the $N!$ permutations of λ . Thus, we are left with minimizing D over the convex polytope defined by the permutations of λ . In minimizing a linear function over a convex polytope, the optimum is always attained at a corner point. This establishes that the optimal transform is a KLT. ■

REFERENCES

- [1] V. K Goyal, J. Zhuang and M. Vetterli, "Transform coding with backward adaptive updates," *IEEE T. Inf. Th.*, July 2000.
- [2] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorizations and Its Applications*, Academic Press, San Diego, 1979.