Theorems and Counterexamples in Transform Coding

Vivek K Goyal
Digital Fountain, Inc.
v.goyal@ieee.org

Abstract— **The best-known results on optimal transforms for transform coding apply only for Gaussian sources coded at high rates. In such situations, the optimal transform is orthogonal and produces independent transform coefficients. Results presented here to some extent reveal the importance of orthogonality and of independent transform coefficients when the source is not necessarily Gaussian and the rate is not necessarily high.**

I. Introduction

Transform coding is a type of source coding characterized by a modular encoder design that includes a linear transformation of the original data and scalar quantization of the resulting coefficients. It arises from applying the "divide and conquer" principle to lossy source coding [3]. As depicted in Fig. 1, a potentially complicated N-dimensional quantizer α is replaced by a linear transform T and N scalar quantizers $\{\alpha_i\}_{i=1}^N$. Constraining the class of source codes like this limits the rate–distortion performance, but also makes design and encoding much simpler.

Transform coding with an orthogonal analysis transform T and perfect reconstruction synthesis transform $U = T^{-1}$ is so common that these constraints are often presupposed in the design of a transform coding system.

Theorem 1 ([1]) Suppose the distortion of each scalar quantizer is given by $E[(y_i - \hat{y}_i)^2] = c\sigma_i^2 2^{-2R_i}$, $i = 1, 2, \ldots, N$, where σ_i^2 is the variance of y_i and R_i is the rate allocated to y_i. (This holds for high-rate coding of a Gaussian source.) An orthogonal transform that minimizes $\prod_{i=1}^N \sigma_i^2$ and hence is optimal when the best (arbitrary real) bit allocation is used is a Karhunen-Loève transform (KLT).

The optimality of KLTs for transform coding of Gaussian sources is actually somewhat more general:

Theorem 2 ([2]) Consider a transform coder with orthogonal analysis transform T and synthesis transform $U = T^{-1} = T^T$. Suppose there is a single function g to describe the quantization of each transform coefficient through $E[(y_i - \hat{y}_i)^2] = \sigma_i^2 g(R_i)$, $i = 1, 2, \ldots, N$. Then for any bit allocation $(R_1, R_2, \ldots, R_N)$ there is a KLT that minimizes the distortion.

II. Theorems

While writing expository works on transform coding [3], [4], I attempted to find more general transform optimality results that do not require (all of the conditions) T orthogonal, $U = T^{-1}$, and x jointly Gaussian. I obtained the following:

Theorem 3 ([4]) In a transform coder with invertible analysis transform T, suppose the transform coefficients are independent. If the component quantizers reconstruct to centroids, T^{-1} is the optimal synthesis transform.

Theorem 4 ([4]) Consider a transform coder in which analysis transform T produces independent transform coefficients, the synthesis transform is T^{-1}, and the component quantizers reconstruct to their respective centroids. To minimize the MSE distortion, it is sufficient to consider transforms with orthogonal rows, i.e., T such that TT^T is a diagonal matrix.

The author's mailing address is Digital Fountain, 39141 Civic Center Drive, Suite 300, Fremont, CA 94538. URL: http://lcavwww.epfl.ch/~goyal/.

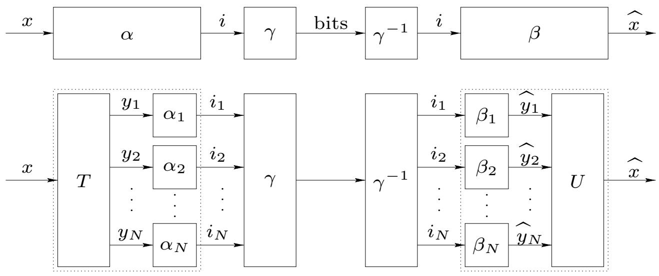

Fig. 1. Any block source code can be decomposed as shown with encoder $\gamma \circ \alpha$ and decoder $\beta \circ \gamma^{-1}$. γ is an entropy code and α and β are the encoder and decoder of an N-dimensional quantizer. In a transform code, α is replaced with a linear transform T and a set of N scalar quantizer encoders. β is replaced with N scalar quantizer decoders and another linear transform U. The intermediate y_is are called transform coefficients. Usually $U = T^{-1}$.

Theorem 5 ([4]) Consider a high-rate transform coding system employing entropy-constrained uniform quantization (ECUQ). A transform with orthogonal rows that produces independent transform coefficients is optimal, when such a transform exists. Furthermore, the norm of the ith row divided by the ith quantizer step size is optimally a constant.

III. Counterexamples

Theorems 3–5 are sharp in the sense that we obtain the following counterexamples. Each will be detailed in the talk.

Example 1: A transform coding system has scalar quantizer decoders $\{\beta_i\}_{i=1}^N$ that reconstruct each quantized transform coefficient to the corresponding centroid. Is $U = T^{-1}$ optimal? No, not necessarily. When transform coefficients are not independent, centroid reconstruction of the scalar components does not imply overall centroid reconstruction. A linear transform $U \neq T^{-1}$ may reduce the distortion. (Consider jointly Gaussian correlated transform coefficients.)

Example 2: T is an orthogonal transform that produces uncorrelated transform coefficients, and ECUQ is used at a high rate with optimal bit allocation. Is T optimal? Not necessarily. A nonorthogonal transform may be superior. (Consider a source uniformly distributed on a parallelogram centered at the origin.)

Example 3: There exists an orthogonal transform T that produces independent transform coefficients, and ECUQ is used with optimal bit allocation. Is T optimal? No, not necessarily when the rate is low. (Consider a source uniformly distributed on a rhombus centered at the origin.)

References

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Acad. Pub., Boston, MA, 1992.

[2] V. K Goyal, J. Zhuang, and M. Vetterli, "Transform coding with backward adaptive updates," *IEEE Trans. Inform. Th.*, vol. 46, no. 4, pp. 1623–1633, July 2000.

[3] V. K Goyal, "Theoretical foundations of transform coding," *IEEE Sig. Proc. Mag.*, vol. 18, no. 5, pp. 9–21, Sept. 2001.

[4] V. K Goyal, *Single and Multiple Description Transform Coding with Bases and Frames*, SIAM, 2002, in preparation.