Performance of Low Rate Entropy-Constrained Scalar Quantizers

Daniel Marco and David L. Neuhoff
Dept. Electrical Engineering and Computer Science
University of Michigan
1301 Beal Ave., Ann Arbor MI 48109, USA
{idaniel, neuhoff}@umich.edu

Abstract — **The operational rate-distortion function of entropy-constrained scalar quantizers in the asymptotic low resolution regime, for memoryless Gaussian sources, is investigated. It is shown that asymptotically, as distortion tends to variance, or equivalently as rate tends to zero, the operational rate-distortion function matches the Shannon rate-distortion function. This implies that scalar quantization is asymptotically optimal, a fact not previously known.**

I. Introduction

Scalar quantizers have mostly been investigated in high resolution, where approximations are more readily available. Here we seek to determine their asymptotic low resolution behavior. In particular we consider variable-rate quantizers, where rate is measured by the quantizers output entropy. (Fixed-rate quantizers are clearly not optimal and hence are not examined.) The distortion measure considered is mean squared error (MSE).

Figure 1 illustrates the operational rate-distortion function and its slope at $D = \sigma^2$, for a Gaussian source with variance σ^2. The dashed line indicates performance in high resolution. The dotted line is a qualitative representation of the operational rate-distortion function in the remaining rate region. The solid line represents the slope of the operational rate-distortion function at $D = \sigma^2$.

In order to show our results, we consider infinite-level uniform threshold scalar quantizers and binary (two-level) scalar quantizers. In both cases the operational rate-distortion functions turn out to be asymptotically the same and match the Shannon rate-distortion function (i.e. their slope at $D = \sigma^2$ is the same). Thus, we also conclude that they asymptotically equal the operational rate-distortion function of scalar quantizers in general. Consequently, the latter asymptotically matches Shannon's rate-distortion function as well.

We mention that in [1] a parametric rate-distortion function was provided for low resolution fixed-rate transform codes, which scalar quantize the transform coefficients. These results are for asymptotically large block lengths.

[1]This work was supported by NSF Grant ANI-0112801.

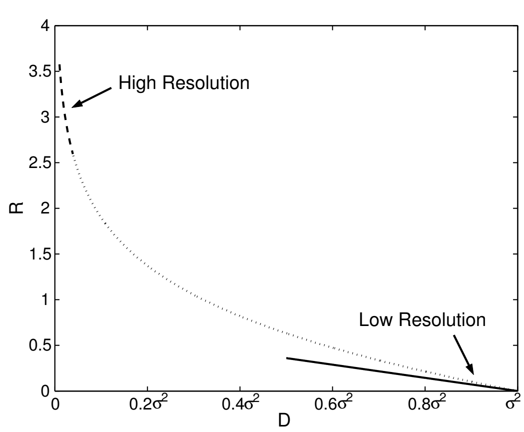

Figure 1: Operational rate-distortion function for a Gaussian source with variance σ^2, and its slope at $D = \sigma^2$.

II. Notation

A uniform quantizer with step size Δ and offset $0 < \alpha < 1$, is a quantizer whose k^{th} cell is $[(k-\alpha)\Delta, (k+1-\alpha)\Delta)$ for all $k \in \mathbb{Z}$. The offset represents the fraction of the $k = 0$ cell that is to the left of the origin. The reconstruction levels are taken to be cell centroids. Let $\lambda = \frac{\Delta}{\sigma}$, where σ^2 is the variance of the zero mean Gaussian source. We let $H_U(\alpha, \lambda)$ denote the entropy of the output of a uniform quantizer with offset α and ratio λ, and $D_U(\alpha, \Delta, \sigma^2)$ denote its MSE.

Similarly, for binary quantizers we let T denote the threshold, so that cell 1 is $(-\infty, T)$ and cell 2 is $[T, \infty)$. The reconstruction levels are centroids. We let $\lambda = \frac{T}{\sigma}$ (this λ will be distinguishable from context from the one defined in the uniform case), and we denote the entropy of a binary quantizer output with ratio λ by $H_B(\lambda)$, and $D_B(T, \sigma^2)$ denotes its MSE. Finally, we let $o_z(1)$ denote a quantity that goes to zero as $z \to \infty$.

III. Results

Theorem 1:

$$H_U(\alpha, \lambda) = \left(\frac{1}{2}\log_2 e\right)\left[\frac{\alpha\lambda}{\sqrt{2\pi}}e^{-\frac{\alpha^2\lambda^2}{2}} + \frac{(1-\alpha)\lambda}{\sqrt{2\pi}}e^{-\frac{(1-\alpha)^2\lambda^2}{2}}\right] \times \left[1 + o_{\alpha\lambda}(1) + o_{(1-\alpha)\lambda}(1)\right] .$$

Theorem 2:

$$1 - \frac{D_U(\alpha, \Delta, \sigma^2)}{\sigma^2} = \left[\frac{\alpha\lambda}{\sqrt{2\pi}}e^{-\frac{\alpha^2\lambda^2}{2}} + \frac{(1-\alpha)\lambda}{\sqrt{2\pi}}e^{-\frac{(1-\alpha)^2\lambda^2}{2}}\right] \times \left[1 + o_{\alpha\lambda}(1) + o_{(1-\alpha)\lambda}(1)\right] .$$

Theorem 3: $H_B(\lambda) = \left(\frac{1}{2}\log_2 e\right)\frac{1}{\sqrt{2\pi}}\lambda e^{-\frac{\lambda^2}{2}}\left[1 + o_\lambda(1)\right]$.

Theorem 4: $1 - \frac{D_B(T,\sigma^2)}{\sigma^2} = \frac{1}{\sqrt{2\pi}}\lambda e^{-\frac{\lambda^2}{2}}\left[1 + o_\lambda(1)\right]$.

Theorem 5: *In the low resolution regime the operational rate-distortion function of uniform, binary and general scalar quantizers, for a memoryless Gaussian source, is given by*

$$\bar{R}_{\sigma^2}(D) = \frac{1}{2}(\log_2 e)[1 - \frac{D}{\sigma^2}][1 + o_{D/\sigma^2}(1)] ,$$

where $o_{D/\sigma^2}(1)$ is a quantity that goes to zero as $\frac{D}{\sigma^2} \to 1$.

References

[1] D.F. Lyons, "Fundamental limits of low-rate transform codes," *Ph.D. Thesis, EECS Dept., University of Michigan*, 1992.