On Optimal Frame Expansions for Multiple Description Quantization

Sanjeev Mehrotra
Stanford Univ./Microsoft Corp.
One Microsoft Way
Redmond, WA 98052
sanjeev@ieee.org

Philip A. Chou
Microsoft Corp.
One Microsoft Way
Redmond, WA 98052
pachou@microsoft.com

Abstract — We study the problem of finding the optimal overcomplete (frame) expansion and bit allocation for multiple description quantization of a Gaussian signal at high rates over a lossy channel.

I. INTRODUCTION

The setup is shown in Figure 1. In multiple description quantization using overcomplete (frame) expansions [1, 2], an input signal \( \mathbf{x} \in \mathbb{R}^N \) is represented by a vector \( \mathbf{y} = \mathbf{F} \mathbf{x} \in \mathbb{R}^N \), \( N > K \). \( \mathbf{F} \) is a \( N \times K \) matrix, called the frame operator. It is assumed any \( K \) rows of \( \mathbf{F} \) span \( \mathbb{R}^N \). The coefficients of \( \mathbf{y} \) are scalar quantized to obtain \( \hat{\mathbf{y}} \) and are then independently entropy coded using on average a total of \( R \) bits allocated among the \( N \) coefficients. In channel state \( s \), the decoder receives \( N_{rs} \leq N \) coefficients after potential erasures, and reconstructs the signal \( \hat{\mathbf{x}} \) from the received coefficients. The number of channel states is \( 2^N \) since each coefficient can be either received or lost. For a given distribution over channel states, we wish to find the frame operator \( \mathbf{F} \) and the bit allocation for the transform coefficients that minimizes the expected squared error \( D = E[\|\mathbf{x} - \hat{\mathbf{x}}\|^2] \) subject to a constraint on the average rate, \( R \), for asymptotically large \( R \) and Gaussian \( \mathbf{x} \).

II. ANALYSIS

Without loss of generality, assume that \( \mathbf{x} \) is distributed with zero mean and diagonal covariance matrix \( \mathbf{R}_\mathbf{x} = \operatorname{diag}(\sigma_0^2, \ldots, \sigma_{N-1}^2) \) (else can use KLT). Let \( \mathbf{q} = \mathbf{y} - \hat{\mathbf{y}} \) be the quantization error and let \( \mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} \) be the reconstruction error. At high rate, assume \( \mathbf{q} \) is distributed with zero mean and diagonal covariance matrix with \( E[\|\mathbf{q}\|^2] = \sigma_q^2 \mathbf{I}^{2\times 2N} \), where \( \sigma_q^2 = \sigma/6 \) if entropy coded uniform scalar quantization is used. The distortion can be written as \( D = \sum_s p_s D_s \), where \( D_s = E[\|\mathbf{e}\|^2 | S = s] \), and \( p_s \) is the probability of the channel being in state \( s \). Let \( \mathbf{y}_{rs} \) denote the \( N_{rs} \) dimensional vector of received coefficients. Let \( \mathbf{F}_{rs} \) be a \( N_{rs} \times K \) matrix consisting of rows of \( \mathbf{F} \) corresponding to the received coefficients.

To obtain an expression for \( D_s \), there are two cases to consider: \( N_{rs} \geq K \) and \( N_{rs} < K \). When \( N_{rs} \geq K \), the decoder has enough information to localize the input vector to a finite cell. Although the actual reconstruction will use a consistent reconstruction [1, 3], for analysis purposes, we use the optimal linear reconstruction as \( \hat{\mathbf{x}} = \mathbf{F}^+ \mathbf{y}_{rs} \), where \( \mathbf{F}^+ \) is the pseudo-inverse of \( \mathbf{F} \). Since \( \mathbf{x} = \mathbf{F}^+ \mathbf{y}_{rs} \), the conditional distortion can be written as \( D_s = E[\|\mathbf{e}\|^2 | S = s] = E[\|\mathbf{F}^+ \mathbf{y}_{rs} - \mathbf{x}\|^2] \). When \( N_{rs} < K \), there is not enough information to localize \( \mathbf{x} \) to a finite cell. In particular \( \mathbf{x} \) is bounded in \( N_{rs} \) dimensions and unbounded in \( K - N_{rs} \) dimensions. Thus, \( \mathbf{x} = \mathbf{F}_{rs} \mathbf{y}_{rs} + (\mathbf{F}_{rs})^\perp \mathbf{y}_{rs} \), where the rows of \( \mathbf{F}_{rs} \) form an orthonormal basis for the subspace orthogonal to the span of the rows of \( \mathbf{F}_{rs} \) and \( \mathbf{y}_{rs} \) is a \( K - N_{rs} \) dimensional vector. Now the optimal linear reconstruction is

\[
\hat{\mathbf{x}} = \mathbf{F}_{rs}^+ \mathbf{y}_{rs} + (\mathbf{F}_{rs})^\perp \mathbf{y}_{rs} + E[\|\mathbf{y}_{rs} - \mathbf{F}_{rs} \mathbf{y}_{rs}\|^2 | \mathbf{y}_{rs}],
\]

which gives a distortion of \( D_s = E[\|\mathbf{F}_{rs}^+ \mathbf{y}_{rs} - \mathbf{x}\|^2] + E[\|\mathbf{y}_{rs} - \mathbf{F}_{rs} \mathbf{y}_{rs}\|^2 | \mathbf{y}_{rs}] = \hat{\mathbf{y}}_s \). Since the source is Gaussian, \( E[\|\mathbf{y}_{rs} - \mathbf{F}_{rs} \mathbf{y}_{rs}\|^2 | \mathbf{y}_{rs}] \) can be easily computed.

Using the equations for \( D_s \) and the fact that \( E[\|\mathbf{q}\|^2] \) is diagonal, the portion of distortion that can be minimized by bit allocation can be written as \( D_b = \sum_{i=0}^{N-1} \alpha_i \sigma_i^2 \), where \( \alpha_i \) is a function of the transform \( \mathbf{F} \), the channel state probabilities \( p_s \), and the quantization constant \( c \). Let \( D_{ab} \) be the remaining portion of the distortion \( D \). Minimizing \( D_b \) is a classic bit allocation problem with solution given by \( p_s = R/N + \log_2 \left( \alpha_i \sigma_i^2 / (\sum_{i=0}^{N-1} \alpha_i \sigma_i^2) \right)^{1/N} / 2 \). This gives an optimal \( D_b \) of \( D_{ab} = N(\sum_{i=0}^{N-1} \alpha_i \sigma_i^2)^{1/N} - 2R/N \). To find the optimal transform, we have to minimize \( D_b + D_{ab} \). Since it is hard theoretically, we use numerical gradient descent techniques by varying one coefficient at a time.

Results show that at high loss rates \( D_b \) is the dominating term which is minimized by repeating the coefficient with highest variance. At low loss rates, \( D_{ab} \) is the dominating term which is minimized by the optimal source code. Results are shown for \( 3 \times 2 \) expansion in Figure 2, where the values for \( \theta_i = \tan^{-1}(F_i/F_0) \), \( i = 0, 1, 2 \) are plotted with rate constraint \( R = 6 \) bits and variances \( \sigma_0^2 = 4 \) and \( \sigma_1^2 = 1 \). Also shown is \( \phi_1 \), which is the kth row of matrix \( \mathbf{F} \).

REFERENCES

