# OPTIMAL BIT ALLOCATION IN THE PRESENCE OF QUANTIZER FEEDBACK

*K. Metin Uz, Jerome M. Shapiro, and Martin Czigler*
The David Sarnoff Research Center,
A Subsidiary of SRI International
CN5300
Princeton, NJ 08543-5300.

**Abstract**

We consider the problem of optimal bit allocation in various forms of predictive coding, where the predictor itself has errors resulting from previous quantization. The solution to this problem has potential application to many forms of image and video coding where predictive coding is employed. In predictive coding, the input to the quantizer can be decomposed into the *innovation*, i.e., the part of the quantizer input that is unpredictable given an unquantized predictor, and the *quantizer feedback*, i.e. the part of the quantizer input signal due to the quantization of the predictor. The natural question that arises is whether it is better to allocate more bits to the predictor since quantization errors persist longer, or to allocate more bits to coding the total residual. This problem is analyzed for predictive video coding through the use of a simple parametric distortion-rate model for the propagation of quantization errors. This model provides a framework in which the optimal bit allocation problem can be solved in the presence of quantizer feedback.

## 1. Introduction

This paper discusses optimal bit allocation in various forms of predictive coding, where the predictor itself has errors resulting from previous quantization. The solution to this problem has potential application to many forms of image and video coding where predictive coding is employed. Two such examples are 1) video coding with temporal prediction where quantization errors on the predictor frame propagate to the predicted frames, and 2) pyramid based image coding, where the quantized coarse level of a pyramid is upsampled and used as a predictor for a higher resolution image. Both of these scenarios are examples of *quantizer feedback*, where quantization errors in the predictor are re-quantized when the prediction error is coded. Since the quantization occurs in a closed-loop, the solution to the optimal bit allocation problem is non-trivial. Allocating more (or less) bits to the predictor can reduce (or increase) the number of bits needed to code the prediction error to a specified quality. The sensitivity of this tradeoff, however, directly depends on the *potential* accuracy of the prediction, which can be measured by the encoder using an unquantized predictor to obtain statistics for bit allocation.

In this paper, optimal bit allocation procedures are derived for a simple parametric model of closed-loop quantizer feedback in the context of predictive video coding. The results for the minimum mean-square-error (MMSE) optimization are an extension of a more traditional open-loop bit allocation procedure often cited in the literature [1]. Although we use video coding as a platform, the results are applicable to any quantizer feedback problem that uses our model. Since the MMSE solution does not produce equal distortion in all frames, we have also included the optimal MINMAX bit allocation that minimizes the frame distortion subject to equal distortions per frame. Although the MMSE solution yields lower average distortion, for predictive video coding, the MINMAX solution may in fact produce a better visual appearance.

Our approach to this problem incorporates the following novel features:

- a simple model consisting of a parametric distortion-rate function in which the quantizer input is explicitly decomposed into 1) the quantization error fed back from the predictor, and 2) the ***innovation*** as defined by the prediction error obtained from using an unquantized predictor;
- an exact MMSE solution that involves solving one non-linear monotonic equation for one Lagrange multiplier, after which the bit allocation has a closed form analytic solution.

## 2. A Parametric Distortion-Rate Model for Predictive Coding

In this section, we consider the situation of predictive video coding for a self-contained sequence of $N$ frames. Frame 1 is encoded spatially, with no quantizer feedback, frames $m = 2, \ldots, N$ are predicted from frame $m-1$, and the prediction residual is encoded spatially. For frame 1,



0-7803-0946-4/93 $3.00 © 1993 IEEE

we assume that the operational distortion-rate function has the exponential form

$$E_1 = e^{-\alpha R_1} X_1, \tag{1}$$

where $E_1$ is the output variance, or the variance of the quantization error, $R_1$ is the rate in bits per pixel, $\alpha$ is the exponential decay parameter of the distortion-rate function, and $X_1$ is a complexity parameter having the same units as the variance.

To assess the utility of this model, consider an input of independent, identically distributed Gaussian pixels. The theoretical distortion-rate function would be identical to (1) by setting $\alpha = 2\ln 2$ and $X_1$ equal to the input variance $\sigma_1^2$ [1]. If the samples were correlated but still Gaussian, the theoretical distortion-rate function would set $X_1$ equal to the theoretical minimum prediction error variance $\eta_1^2 = \gamma_1 \sigma_1^2$, where $\gamma_1$ is the spectral flatness measure [1]. Since practical situations use real coders, and non-gaussian inputs, the model is sufficiently general to handle other exponential decay parameters and any multiplicative constant to the input variance. These model parameters, $\alpha$ and $X_1$, can be estimated by operating a spatial encoder at several different rates and measuring the resulting distortion. For example, after taking logarithms in (1), the parameters can be estimated using simple least-squares.

Assume now that frame $m$ is predicted from frame $m-1$. A reasonable model for the distortion rate function is given by

$$E_m = e^{-\alpha R_m}(X_m + \rho_m E_{m-1}), \tag{2}$$

where $E_{m-1}$ is the output variance of frame $m-1$, and $\rho_m : 0 \leq \rho_m \leq 1$ is the *coefficient of quantizer feedback* and represents the fraction of the error of the previous frame that is actually used in the prediction. $X_m$ is the complexity parameter of the *innovation*, i.e., the residual that would result if frame $m$ was predicted from an unquantized frame $m-1$. Equation (2) also applies to the spatially encoded frame by setting $\rho_m = 0$.

Note that for the following solution to apply, the exponential decay parameter $\alpha$, must be the same for all frames. Thus, to estimate the model parameters for an sequence of $N$ frames, there are $N+1$ unknown parameters: $\alpha$ and $X_i, i = 1, \ldots, N$. In this case, to gather statistics, a spatial encoder must be run at least twice on the innovation for each frame.

## 3. MMSE Optimal Bit Allocation

In this section, a solution procedure for the optimal MMSE bit allocation is developed using Lagrange multipliers. Given frames $m = 1, \ldots, N$ with known statistics $X_m$, we want to choose the bit allocation, $R_1, \ldots, R_n$ to minimize the total (or average) MSE distortion

$$D = \sum_{m=1}^{N} E_m, \tag{3}$$

subject to a budget constraint on the total number of bits

$$\sum_{m=1}^{N} R_m = NR. \tag{4}$$

It is also assumed that the rates must all be nonnegative. In the event that the solution yields negative rates, an active set method can be used to reduce the number of unknowns in the problem so that all the rates are nonnegative. For those frames for which the rate is set to zero, the quantizer feedback must still be propagated in the revised formulation.

The Lagrangian is given by

$$\mathcal{L} = D + \alpha\lambda\left(\sum_{m=1}^{N} R_m - NR\right), \tag{5}$$

where the scale factor in the Lagrange multiplier will make the notation simpler in the solution.

The pair of equations given by (1) and (2) provide a set of one-step prediction relationships between the quantization errors of adjacent frames. It is also useful to derive a higher-order forward prediction equation. By expanding (1) and (2), we can explicitly express $E_n$ in terms of the complexities $X_1, \ldots, X_n$

$$E_n = e^{-\alpha R_n} X_n + \sum_{i=1}^{n-1} e^{-\alpha \sum_{j=i}^{n} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i. \tag{6}$$

We can express $E_n$ in terms of $E_m$ for $n > m$ by separating contributions to $E_n$ from frames $1, \ldots, m$ and from frames $m+1, \ldots, n$ in (6) as follows:

$$\begin{aligned} E_n &= e^{-\alpha R_n} X_n \\ &\quad + \sum_{i=m+1}^{n-1} e^{-\alpha \sum_{j=i}^{n} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i \\ &\quad + \sum_{i=1}^{m} e^{-\alpha \sum_{j=i}^{n} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i. \end{aligned} \tag{7}$$

Factoring out the exponential factors and the $\rho_i$ that depend on frames $i > m$ from the last term yields

$$\begin{aligned} E_n &= e^{-\alpha R_n} X_n + \sum_{i=m+1}^{n-1} e^{-\alpha \sum_{j=i}^{n} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i \\ &\quad + e^{-\alpha \sum_{j=m+1}^{n} R_j} \left(\prod_{k=m+1}^{n} \rho_k\right) \\ &\quad \left[e^{-\alpha R_m} X_m + \sum_{i=1}^{m-1} e^{-\alpha \sum_{j=i}^{m} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i\right] \end{aligned} \tag{8}$$

$$= e^{-\alpha R_n} X_n + \sum_{i=m+1}^{n-1} e^{-\alpha \sum_{j=i}^{n} R_j} \left(\prod_{k=i+1}^{n} \rho_k\right) X_i$$

$$+e^{-\alpha\sum_{j=m+1}^{n} R_j}(\prod_{k=m+1}^{n} \rho_k)E_m. \qquad (9)$$

Notice that the dependence of $E_n$ on $R_m$ is expressed solely in the last term of (9) through $E_m$.

To obtain a set of optimality equations, the first derivative of the Lagrangian is set to zero. For optimality, the marginal error due to the bit allocation at each frame is given by

$$\frac{\partial D}{\partial R_m} = -\alpha\lambda. \qquad (10)$$

From (3) and (9), we see that

$$\begin{aligned} \frac{\partial D}{\partial R_m} &= \sum_{n=m}^{N} \frac{\partial E_n}{\partial R_m} && (11)\\ &= [\sum_{n=m+1}^{N} e^{-\alpha\sum_{j=m+1}^{n} R_j}(\prod_{k=m+1}^{n} \rho_k)+1]\frac{\partial E_m}{\partial R_m} && (12)\\ &= -\alpha[K_m+1]E_m, && (13) \end{aligned}$$

where

$$K_m \triangleq \sum_{n=m+1}^{N} e^{-\alpha\sum_{j=m+1}^{n} R_j}(\prod_{k=m+1}^{n} \rho_k), \qquad (14)$$

and we have made use of the fact from (2) that

$$\frac{\partial E_m}{\partial R_m} = -\alpha E_m. \qquad (15)$$

Note that for the marginal error associated with any $R_m$ there is a component of marginal error associated with the effect of $R_m$ directly on $E_m$ and a component of marginal error associated with the effect of $E_m$ on future frames. Therefore we can combine (10) and (13) to obtain

$$[K_m+1]E_m = \lambda, \qquad (16)$$

where $K_m$ is referred to as the *feedback component of marginal error*. From (14), we can derive a one-step backward prediction for $K_m$

$$K_m = e^{-\alpha R_{m+1}}\rho_{m+1}[K_{m+1}+1], \quad K_N = 0. \qquad (17)$$

Using (16) to substitute in for $[K_{m+1}+1]$ results in

$$K_m = e^{-\alpha R_{m+1}}\rho_{m+1}\frac{\lambda}{E_{m+1}}. \qquad (18)$$

Using the one-step ahead predictor (2) to solve for $E_{m+1}$ yields

$$K_m = \frac{\lambda\rho_{m+1}}{X_{m+1}+\rho_{m+1}E_m}. \qquad (19)$$

Substituting (19) into (16) and expanding, we obtain a quadratic equation for $E_m$:

$$\rho_{m+1}E_m^2 + X_{m+1}E_m - \lambda X_{m+1} = 0, \qquad (20)$$

which provides a solution to the minimum error variance for each frame under optimality:

$$E_m = \frac{\lambda}{(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}+\frac{1}{2}}. \qquad (21)$$

Note that in the absence of quantizer feedback, i.e. when $\rho_m = 0$ for all frames, the optimal solution yields equal distortion for each frame. However in the case of quantizer feedback, frames that are useful as predictors, i.e. small $X_{m+1}$ are encoded with lower distortion then frames that are poor predictors, i.e. large $X_{m+1}$, or $\rho_{m+1} = 0$ in the case when the next frame is encoded spatially.

Using (21) to substitute for both $E_{m-1}$ and $E_m$ in (2) and solving for $e^{-\alpha R_m}$ yields

$$e^{-\alpha R_m} = \frac{\lambda}{X_m[(\frac{1}{4}+\frac{\lambda\rho_m}{X_m})^{\frac{1}{2}}+\frac{1}{2}][(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}+\frac{1}{2}]}. \qquad (22)$$

Taking logarithms, we obtain the optimal bit allocation in terms of the known statistics and the Lagrange multiplier

$$\begin{aligned} R_m &= \frac{1}{\alpha}\ln X_m + \frac{1}{\alpha}\ln[(\frac{1}{4}+\frac{\lambda\rho_m}{X_m})^{\frac{1}{2}}+\frac{1}{2}] \\ &\quad +\frac{1}{\alpha}\ln[(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}+\frac{1}{2}] - \frac{1}{\alpha}\ln\lambda. \qquad (23) \end{aligned}$$

To complete the solution, we apply the constraint on the rate (4) which yields the non-linear equation

$$R = \frac{1}{\alpha N}\sum_{m=1}^{N}\ln X_m - g(\lambda), \qquad (24)$$

where

$$\begin{aligned} g(\lambda) &= \frac{1}{\alpha}\ln\lambda - \frac{1}{\alpha N}\sum_{m=1}^{N}\ln[(\frac{1}{4}+\frac{\lambda\rho_m}{X_m})^{\frac{1}{2}}+\frac{1}{2}] \\ &\quad -\frac{1}{\alpha N}\sum_{m=1}^{N}\ln[(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}+\frac{1}{2}] \qquad (25) \end{aligned}$$

To help characterize the solution, we want to examine $g(\lambda)$. Differentiating and simplifying, we have

$$\begin{aligned} \frac{dg(\lambda)}{d\lambda} &= \frac{1}{4\alpha N\lambda}\sum_{m=1}^{N}(\frac{1}{(\frac{1}{4}+\frac{\lambda\rho_m}{X_m})^{\frac{1}{2}}}+\frac{1}{(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}}) \\ &> 0, \qquad (26) \end{aligned}$$

indicating that $g(\lambda)$ is monotonically increasing. Therefore (25) has a unique solution for $\lambda$.

Given the $X_i$ and $\alpha$, a practical solution for the optimal bit allocation requires solving (25) for $\lambda$ by Newton's method or successive-approximation, and then solving (23) directly.

For comparison with the solution obtained in the absence of quantizer feedback, equations (23), (24), and (25) can be combined:

$$R_m = R + \frac{1}{\alpha}\ln X_m - \frac{1}{\alpha N}\sum_{n=1}^{N}\ln X_n + b_m + f_m, \quad (27)$$

where

$$\begin{aligned} b_m &= +\frac{1}{\alpha}\ln[(\frac{1}{4}+\frac{\lambda\rho_m}{X_m})^{\frac{1}{2}}+\frac{1}{2}] \\ &\quad -\frac{1}{\alpha N}\sum_{n=1}^{N}\ln[(\frac{1}{4}+\frac{\lambda\rho_n}{X_n})^{\frac{1}{2}}+\frac{1}{2}], \quad (28) \\ f_m &= +\frac{1}{\alpha}\ln[(\frac{1}{4}+\frac{\lambda\rho_{m+1}}{X_{m+1}})^{\frac{1}{2}}+\frac{1}{2}] \\ &\quad -\frac{1}{\alpha N}\sum_{n=1}^{N}\ln[(\frac{1}{4}+\frac{\lambda\rho_{n+1}}{X_{n+1}})^{\frac{1}{2}}+\frac{1}{2}]. \quad (29) \end{aligned}$$

Note that the first three terms in (27) represents the optimal bit allocation in the absence of quantizer feedback. The other terms represent adjustments to allocation that depend on whether frame $m$ is predicted, used as a predictor, or both. The term $b_m$ represents the necessary *backward* adjustment to $R_m$ that results from predicting frame $m$ from frame $m-1$, whereas, $f_m$ represents the necessary *forward* adjustment to $R_m$ that results from using frame $m$ as a predictor for frame $m+1$. The forward adjustment $f_m$, which causes an upward adjustment in the bit allocation $R_m$ when frame $m$ is a good predictor for frame $m+1$ occurs because reducing the quantization error in frame $m$ also reduces the input variance to the quantizer for frame $m+1$. Therefore for the same $X_m$, frame $m$ should get more bits if it is a good predictor than it would get if it were a poor predictor. The backward adjustment, which causes an upward adjustment in the bit allocation $R_m$ when frame $m$ is well predicted from frame $m-1$. In that case, $X_m$ is not an accurate measure of the true cost of encoding the prediction residual since a larger than average fraction of the quantizer input variance is due to quantization error from frame $m-1$. In this case, a small value of $X_m$ should not cause a misleadingly low bit allocation.

## 4. Alternative Solution: MINMAX

One of the drawbacks of the MMSE solution is that a frame that is not used as a predictor is given the largest error. In certain situations, this may be an undesirable result. For example, in predictive video coding, the optimal MMSE solution would force the last frame, which is not used for prediction, to have the largest error.

It is also possible to formulate the optimization problem to minimize the maximum (MINMAX) frame distortion. In the absence of non-negativity constraints on the rates, the solution yields equal distortions for each frame. Setting the error variance for each frame equal $E_i = E$, the problem is to

minimize $E$ subject to (4)

The solution procedure is to solve (2) for $R_i$

$$R_i = \frac{1}{\alpha}\ln\frac{X_m + \rho_m E}{E}. \quad (30)$$

Applying (4) leads to a polynomial equation for $E$:

$$E^N e^{N\alpha R} - \prod_{m=1}^{N}(X_m + \rho_m E) = 0. \quad (31)$$

Since $\rho_1 = 0$ and the polynomial equation has only one sign change, there is only one positive root and therefore (31) can be solved uniquely for $E$. Once $E$ is known, (30) yields the optimal MINMAX bit allocation.

## 5. Significance of the Results

This work provides a methodology for performing optimal bit allocation in the presence of quantizer feedback. We also relate the closed-loop bit allocation to the more standard open-loop bit allocation procedure. Although our development is primarily theoretical, the results provide insight and have implications for practical bit allocation problems such as those found in many areas of image and video coding.

An interesting, although perhaps counterintuitive result of this work, is that in the case of temporal prediction of video, the optimal MMSE bit allocation does not yield equal quality for each frame as is the case with no quantizer feedback. In the case of quantizer feedback, frames that are either easily predicted, or good predictors are encoded to higher quality since propagating quantization errors contribute to the total error in predicted frames.

## 6. References


[1] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall: Englewood Cliffs, NJ 1984.