ℓ_2 Optimized Predictive Image Coding with ℓ_∞ Bound

ℓ_2 OPTIMIZED PREDICTIVE IMAGE CODING WITH ℓ_∞ BOUND

BY

SCEUCHIN CHUAH, B.Eng.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2012) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	ℓ_2 Optimized Predictive Image Coding with ℓ_∞ Bound
AUTHOR:	Sceuchin Chuah B.Eng., (Electrical Engineering) McMaster University, Hamilton, Ontario, Canada
SUPERVISOR:	Dr. Xiaolin Wu
NUMBER OF PAGES:	xi, 52

To My Mother and Father

Abstract

In many scientific, medical and defense applications of image/video compression, an ℓ_∞ error bound is required. However, pure ℓ_∞-optimized image coding, colloquially known as near-lossless image coding, is prone to structured errors such as contours and speckles if the bit rate is not sufficiently high; moreover, previous ℓ_∞-based image coding methods suffer from poor rate control. In contrast, the ℓ_2 error metric aims for average fidelity and hence preserves the subtlety of smooth waveforms better than the ℓ_∞ error metric and it offers fine granularity in rate control; but pure ℓ_2-based image coding methods (e.g., JPEG 2000) cannot bound individual errors as ℓ_∞-based methods can. This thesis presents a new compression approach to retain the benefits and circumvent the pitfalls of the two error metrics.

Acknowledgements

I would like express my sincere thanks to Dr. Xiaolin Wu and Dr. Sorina Dumitrescu for their invaluable supervision, guidance and support. I cannot thank them enough for all the insightful critique and feedback that they have given me. Their enthusiasm and high standards in research are indeed admirable and have constantly motivated me to do better.

Special thanks goes to my defence committee members for their helpful reviews and comments. I am also very grateful to all the staff members in the Electrical & Computer Engineering department, especially Cheryl Gies for all her kind assistance, and Dr. Nicola Nicolici for being an inspiration to me with his strong passion in engineering and teaching. Sincere thanks to my fellow colleagues, classmates and friends, Jiantao, Yong, Ying, Zahra, Cong, Maggie, Taimur, Kelvin, Matthew, Shuwen, Dragan, Emily, Karyn, Sagar, Jason, Daniel, Samantha, and Parathy. Their friendship and help has made my master's study very special and memorable.

Last but not least, my heartfelt gratitude goes to my loving parents, Lian Siah Chuah and Sook Yi Mun, for their continuous motivation and support in everything that I do.

Abbreviations

ACL	Asymptotic Closed-Loop
BPP	Bits Per Pixel
CALIC	Context Adaptive Lossless Image Compression
CL	Closed-Loop
DPCM	Differential Pulse Code Modulation
ECSQ	Entropy-Constrained Scalar Quantizer
GAP	Gradient-Adjusted Prediction
JPEG	Joint Photographic Experts Group
MSE	Mean Squared Error
OL	Open-Loop
PMF	Probability Mass Function
PSNR	Peak Signal-to-Noise Ratio
WDAG	Weighted Directed Acyclic Graph

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Problem Statement

There are many important image compression applications in science, medicine, space exploration, precision engineering, etc., where high fidelity of image reconstruction is a crucial requirement. The ideal solution is lossless compression. However, this solution demands a relatively high bit budget, since in spite of intense research in this area, lossless compression rates have remained high (typically 3-4 bpp). The alternative solution for reducing the bit rate is to accept some small loss, and this realization led researchers to develop what became known as near-lossless compression algorithms. Since such algorithms target applications where maintaining small image details in the reconstruction is important, researchers have adopted, as a near-lossless compression criterion, the requirement that a predefined upper bound on the reconstruction error for each pixel be obeyed. In other words, near-lossless image compression became synonymous with ℓ_∞-constrained image compression.

An obvious and simple way to achieve ℓ_∞-constrained image compression is by pre-quantizing first the pixel values via a uniform scalar quantizer, and then to losslessly encode the pre-quantized image (Zandi *et al.*, 1995; Said and Pearlman, 1996). To achieve an ℓ_∞ bound equal to some value τ, a step size of $2\tau + 1$ is used in the scalar quantizer. More efficient near-lossless coding algorithms however, utilize predictive coding with the quantization of residuals (Chen and Ramabadran, 1994; Ke and Marcellin, 1998; Wu *et al.*, 1995; Wu and Bao, 2000; Wu *et al.*, 2011). The algorithm in (Chen and Ramabadran, 1994) is a DPCM coding technique which employs context-based source modeling and arithmetic coding for lossless compression of quantized prediction errors. In order to achieve an ℓ_∞ bound of $\tau = 1$, two different scalar midtread quantizers were suggested in (Chen and Ramabadran, 1994). One of them is uniform with all step sizes equal to 3. The other one is nearly uniform with a bin of size 3 around 0 and all the other bins of size 2. The system in (Ke and Marcellin, 1998) is also based on the DPCM method, but it incorporates an additional mechanism to minimize the entropy of the sequence of quantized prediction residues using a so-called DPCM-trellis. The trellis state transitions restrict the possible pixel reconstructions to those within a τ-error bound. An iterative algorithm determines the trellis path corresponding to the minimum entropy sequence of quantized residuals for each image row. Finally, ℓ_∞-constrained (or near-lossless) CALIC (Wu *et al.*, 1995), which is a variant of lossless CALIC (Wu and Memon, 1997), incorporates a uniform scalar quantizer for the residual errors in the prediction loop. In particular, a step size of $(2\tau + 1)$ is used for the quantizer in order to ensure an error bound no larger than τ.

Among the aforementioned ℓ_∞-constrained image coders, near-lossless CALIC

achieves the highest compression performance when $\tau \leq 3$ (Wu *et al.*, 1995). Further enhancements of near-lossless CALIC were proposed in (Wu and Bao, 2000; Wu *et al.*, 2011) which led to superior performance in terms of bit rate and/or ℓ_2 distortion. However, these techniques incur increased computational complexity either at the encoder, in (Wu and Bao, 2000), where adaptive context modeling is used, or at the decoder, in (Wu *et al.*, 2011), where the hard decision decoding is followed by an ℓ_2 image restoration step.

As seen from the above discussion existing ℓ_∞-constrained image coders generally aim at reaching the smallest bit rate for each given error bound τ and disregard other fidelity criteria. Thus, the number of achievable bit rates is small, only equal to the number of possible values of small τ. Such a coarse rate granularity makes it impossible to finely control the bit rate and image quality. Additionally, the ℓ_∞ constraint alone might not be sufficient to guarantee the necessary level of quality. In particular, ℓ_∞-constrained image coders may introduce artifacts in the form of structured errors even when the value of τ is as low as 4. Fig. 1.1 illustrates such an example with comparisons between ℓ_∞-based image coding (near-lossless CALIC) and ℓ_2-based image coding (JPEG 2000) when compression rates are the same. ℓ_∞-constrained CALIC is seen to create unnecessary speckles and contours in Fig. 1.1c. Though ℓ_2-based JPEG 2000 did not produce those artifacts, it smoothed out the image by too much, removing fine details and distorting edges in Fig. 1.1b.

As seen in the above example the ℓ_2 distortion measure targets for average fidelity and therefore preserves better the smoothness of waveforms. We conclude that it is desirable to include the ℓ_2 criterion in the code design. However, decreasing the ℓ_2 error while maintaining the same bit-rate and error bound might be difficult

Table 1.1: Achievable bit rates and PSNR values for image 4.5e using Near-lossless CALIC

τ	Bit Rate (bpp)	PSNR (dB)
0	3.53	∞
1	2.13	49.95
2	1.57	45.38
3	1.27	42.53
4	1.07	40.47
5	0.94	38.84
6	0.82	37.43
7	0.73	36.30
8	0.67	35.25

without increasing the coding complexity. In contrast, a coder with fine granularity offers a trade-off between bit-rate and fidelity, making it possible to achieve a higher fidelity at the expense of a small increase in the bit rate. Unfortunately, good existing ℓ_∞-constrained image coders lack this ability. To exemplify, consider Table 1.1 which presents the achievable bit rates for the test image in Fig. 4.5e encoded with near-lossless CALIC at $\tau = 0, 1, 2, \ldots, 8$, where the case $\tau = 0$ corresponds to lossless compression using CALIC. Notice the big gaps between consecutive bit rates, especially as τ decreases, which is the case of interest. In order to improve the reconstruction quality at $\tau = 2$, the next available option, which is $\tau = 1$, needs a significant increase in bit rate of 0.56 bpp, while we would like to be able to do this with only a small increase in bit rate. Therefore, we conclude that the ability to achieve a finer rate granularity and the incorporation of the ℓ_2 metric in the design would be beneficial features for ℓ_∞-constrained image coders.

In this thesis we propose an ℓ_2-optimized ℓ_∞-constrained image coding algorithm with fine rate granularity. The proposed coder is a modification of near-lossless CALIC, which replaces the uniform scalar quantization of prediction errors

by context-based ℓ_2-optimized quantization. Specifically, the scalar quantizers for the different contexts are optimized to minimize the ℓ_2 distortion while obeying a constraint on the average entropy over all quantizers and a specified ℓ_∞ error bound.

Optimization of the quantizer used in the prediction loop has been proposed in the past in (Max, 1960; Netravali, 1977; Sharma, 1978) for the scalar case and in (Cuperman and Gersho, 1985; Khalil *et al.*, 2001; Khalil and Rose, 2003) for the vector case. One difficulty encountered when addressing this problem resides in obtaining a distribution that accurately represents the distribution of the prediction errors. This is because the statistics of the residuals depends on the quantizer. Most authors have generated the training set of prediction errors by using the unquantized pixel values in the prediction. This method is known as the open-loop (OL) approach. To address the statistical mismatch of OL, the closed-loop (CL) and the asymptotic closed-loop (ACL) approaches were proposed in (Cuperman and Gersho, 1985) and (Khalil *et al.*, 2001), respectively. They performed the design process iteratively, the quantizer optimized at each iteration being used to obtain the training set of residuals for the next iteration.

In this thesis we adopt the OL approach. The reasoning behind this choice is that, in near-lossless compression, small values of τ are of interest, and in this case, the OL approach provides a good enough approximation of the true statistics of prediction errors. We point out that in this thesis, after collecting the statistics of residuals for each context from a training set of images, each conditional distribution is approximated by a Laplacian distribution which is further used in the optimization.

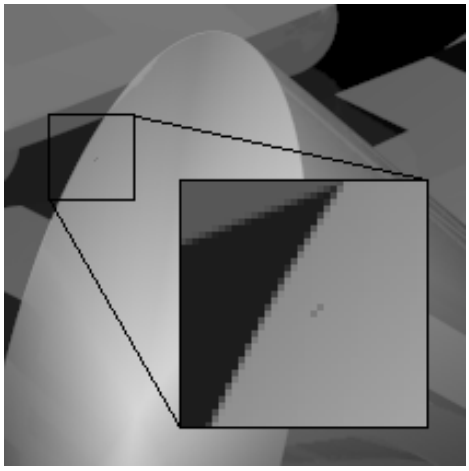
What distinguishes our work from previous work on optimal quantizer design is mainly the criterion used in the optimization. Most quantizer design algorithms aim

at minimizing the ℓ_2 distortion for fixed number of quantizer levels, or for an entropy constraint. We are not aware of any work which incorporates the ℓ_∞ constraint alongside. Scalar quantizer design algorithms mainly fall into one of the following categories: 1) Lloyd-Max method (Max, 1960; Lloyd, 1982), which iteratively optimizes the encoder and the decoder respectively while keeping the other component fixed, and 2) combinatorial algorithms (Bruce, 1964; Sharma, 1978; Wu, 1991; Wu and Zhang, 1993; Dumitrescu and Wu, 2004, 2005, 2007; Muresan and Effros, 2008). While the first approach ensures only a locally optimal solution, the latter algorithms guarantees global optimality when the source alphabet is finite. Our optimization problem requires simultaneous optimization of all quantizers corresponding to different contexts under a common constraint on the average entropy and the ℓ_∞ error bound. We first convert the problem to a Lagrangian formulation as is the common practice in entropy-constrained quantizer design. Interestingly, the Lagrangian formulation allows for separate optimization of each quantizer. We further show that the latter problem can be modeled as a minimum weight path problem. This model is similar in spirit to that used in (Muresan and Effros, 2008). However, we emphasize that while in (Muresan and Effros, 2008) only the ℓ_2-optimization of entropy-constrained quantizers were considered, our problem is more complex by additionally imposing an ℓ_∞ constraint.

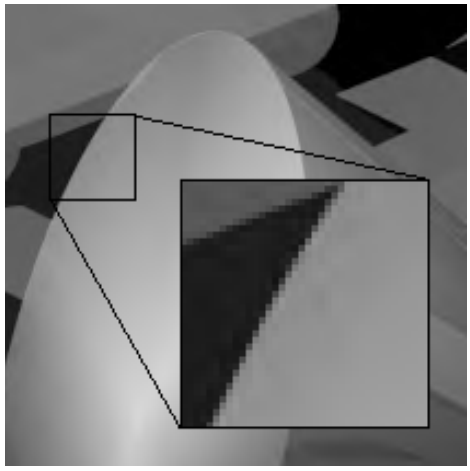
The proposed image coder is able to achieve a much denser set of bit rates than near-lossless CALIC by using different values for the Lagrangian multiplier in the cost function. As our experiments performed on images outside the training set show, the ℓ_∞ constraint enforced in our algorithm allows us to achieve ℓ_∞ error bounds that are always lower than those of JPEG 2000. Meanwhile, the minimization of the ℓ_2

distortion incorporated in the design leads to better ℓ_2 performance than JPEG 2000 above a certain threshold bit rate for each image, a threshold which can be as low as 1.1 bpp. Furthermore, the fine granularity allows for the reconstruction quality to be improved by having only small additions to the used bit rate. In particular, it can be seen in Fig. 1.1d that the proposed coder eliminates the artifacts observed in Fig. 1.1c at the expense of only a very small increase in the bit rate.

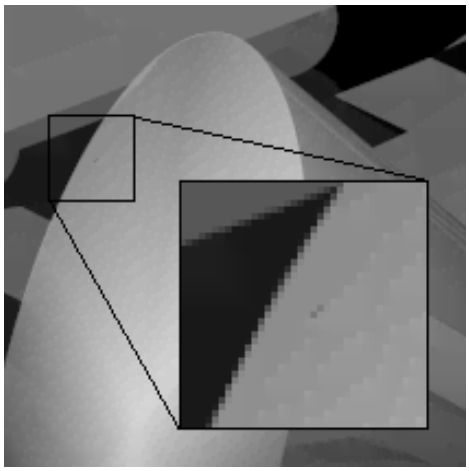
The remainder of this thesis is organized as follows. In Chapter 2 we briefly describe how near-lossless CALIC operates. Then in Chapter 3, we formulate the problem of ℓ_2 optimization of the different context-based quantizers with a common constraint on the average entropy and an ℓ_∞ error bound. Further, we describe the solution algorithm based on the graph approach. We subsequently integrate the optimal scalar quantizers obtained in Chapter 3 into near-lossless CALIC and present extensive experimental results in Chapter 4. The results include performance comparisons with JPEG 2000 and near-lossless CALIC. Finally, conclusions are given in Chapter 5.



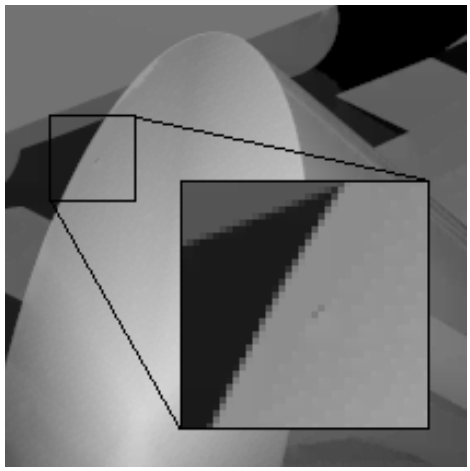
(a) Original image



(b) JPEG 2000



(c) Near-lossless CALIC



(d) Proposed

Figure 1.1: Comparison using a computer-generated test image between (a) Original image / Lossless image coding (Lossless CALIC at rate 1.27 bpp); (b) ℓ_2-based image coding (JPEG 2000 at rate 0.38 bpp, PSNR 42.08dB, ℓ_∞ error bound 22) with details blurred out; (c) ℓ_∞-based image coding (Near-lossless CALIC at rate 0.38 bpp, PSNR 39.45dB, ℓ_∞ error bound 4) with speckles and contours as artifacts; and (d) Proposed method (at rate 0.47 bpp, PSNR 39.63dB, ℓ_∞ error bound 4) preserving details with minimal artifacts.

Chapter 2

Near-lossless CALIC

As shown in the flow diagram in Fig. 2.1, l_∞-constrained CALIC in (Wu *et al.*, 1995) consists of five main components: gradient-adjusted prediction (GAP), uniform quantization, context formation and quantization, context modeling, and entropy coding. We will only briefly describe the encoder since the decoder is just the encoder process reversed.

Let I be the current pixel value to be encoded. The GAP module makes a prediction $\bar{I}$ of I based on the knowledge of the reconstructed pixels $\tilde{I}$ in a precisely defined neighbourhood as shown in Fig. 2.2. We first estimate the horizontal gradient d_h and vertical gradient d_v of the intensity function at current pixel I by the following:

$$\begin{aligned} d_h &= |\tilde{I}_w - \tilde{I}_{ww}| + |\tilde{I}_n - \tilde{I}_{nw}| + |\tilde{I}_n - \tilde{I}_{ne}|, \\ d_v &= |\tilde{I}_w - \tilde{I}_{nw}| + |\tilde{I}_n - \tilde{I}_{nn}| + |\tilde{I}_{ne} - \tilde{I}_{nne}|. \end{aligned} \tag{2.1}$$

Then, the GAP module constructs the predictor $\bar{I}$ by comparing d_h and d_v with the

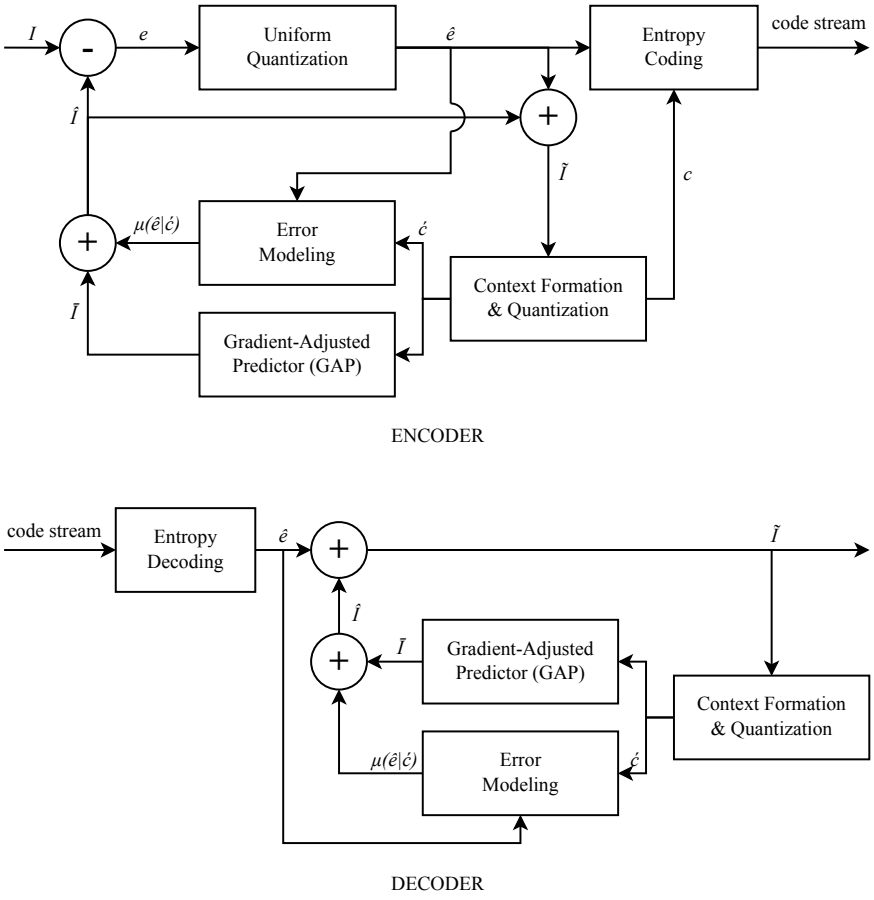


Figure 2.1: Schematic description of near-lossless CALIC.

following procedure:

IF $(d_v - d_h > 80)$ {sharp horizontal edge} $\quad \bar{I} = \tilde{I}_w$

ELSE IF $(d_v - d_h < -80)$ {sharp vertical edge} $\quad \bar{I} = \tilde{I}_n$

ELSE {

$\bar{I} = (\tilde{I}_w + \tilde{I}_n)/2 + (\tilde{I}_{ne} - \tilde{I}_{nw})/4$

IF $(d_v - d_h > 32)$ {horizontal edge} $\quad \bar{I} = (\bar{I} + \tilde{I}_w)/2$

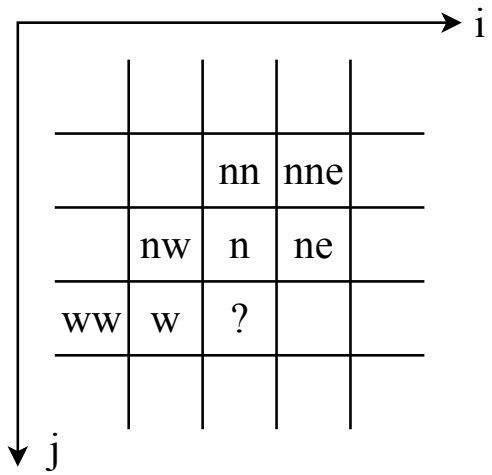


Figure 2.2: Labeling of neighbouring pixels used in prediction and modeling.

$$\begin{aligned}
&\text{ELSE IF } (d_v - d_h > 8) \text{ \{weak horizontal edge\}} \quad \bar{I} = (3\bar{I} + \tilde{I}_w)/4 \\
&\text{ELSE IF } (d_v - d_h < -32) \text{ \{vertical edge\}} \quad \bar{I} = (\bar{I} + \tilde{I}_n)/2 \\
&\text{ELSE IF } (d_v - d_h < -8) \text{ \{weak vertical edge\}} \quad \bar{I} = (3\bar{I} + \tilde{I}_n)/4 \\
\}
\end{aligned} \tag{2.2}$$

The prediction $\bar{I}$ is further improved to $\hat{I}$ by adding the conditional sample mean of the quantized prediction errors $\mu(\hat{e}|\acute{c})$ conditioned on the error modeling context $\acute{c}$. The number of error modeling contexts $\acute{c}$ considered in CALIC is 576 or higher and they are formed based on both the energy level and image texture. The resulting prediction error (or residue) $e = I - \hat{I}$ is then quantized with a uniform scalar quantizer generating $\hat{e}$. The reconstructed pixel value $\tilde{I} = \hat{I} + \hat{e}$ and the quantized prediction residue $\hat{e}$ are fed back into the system to be used in the encoding of future pixels in the image.

Finally, the sequence of quantized prediction residues is losslessly entropy coded by means of a context-based arithmetic coder. Only eight coding contexts $c \in \{1, 2, \ldots, 8\}$ are used for this purpose. They are formed by quantizing an error energy estimator Δ into eight bins based on a criterion to minimize the conditional entropy of prediction errors $-\sum_e p(e) \log p(e|c)$, where $p(e)$ and $p(e|c)$ are respectively the probability and conditional probability of prediction residue e when a uniform quantizer of step size $2\tau + 1$ is used. Δ is more specifically the weighted sum of gradients of the reconstructed pixel values in the neighbourhood and is computed as follows:

$$\Delta = d_h + d_v + |\hat{e}_w|. \tag{2.3}$$

The quantization intervals used for quantizing Δ are shown in Table 2.1 below.

Table 2.1: Quantization intervals for error energy estimator Δ to form eight coding contexts c

τ	c							
	1	2	3	4	5	6	7	8
0	[0,7]	(7,17]	(17,28]	(28,46]	(46,65]	(65,91]	(91,148]	>148
1	[0,2]	(2,6]	(6,11]	(11,23]	(23,47]	(47,72]	(72,140]	>140
2	[0,3]	(3,6]	(6,15]	(15,30]	(30,53]	(53,81]	(81,159]	>159
3	[0,2]	(2,5]	(5,21]	(21,45]	(45,67]	(67,116]	(116,300]	>300
4	[0,4]	(4,13]	(13,39]	(39,68]	(68,94]	(94,127]	(127,165]	>165
5	[0,6]	(6,62]	(62,89]	(89,124]	(125,172]	(173,230]	(230,300]	>300
6	[0,5]	(5,18]	(18,56]	(56,98]	(98,138]	(138,184]	(184,219]	>219
7	[0,4]	(4,18]	(18,53]	(53,89]	(89,124]	(124,183]	(183,300]	>300
8	[0,2]	[0,13]	(13,47]	(47,100]	(100,140]	(140,188]	(188,300]	>300

Chapter 3

Optimal Context-based Quantization of Prediction Errors

The proposed image coder shown in Fig. 3.1 replaces the uniform quantizer of prediction errors in near-lossless CALIC by scalar quantizers optimized for each individual coding context. The optimization problem aims at minimizing the average ℓ_2 distortion over all eight coding contexts while preserving a maximum error bound for each prediction error and a limit on the average output entropy over all eight quantizers. The optimization is performed assuming known distributions of prediction errors and coding contexts.

In this chapter we will address this optimization problem. Section 3.1 presents the definitions and notations pertinent to ℓ_∞-constrained scalar quantizers. Section 3.2 introduces the mathematical statement of the problem and converts it to an unconstrained Lagrangian minimization problem which, it turns out, can be carried out by separately optimizing the quantizer for each coding context. Section 3.3 presents the solution to the latter problem by modeling it as a minimum weight path problem in a

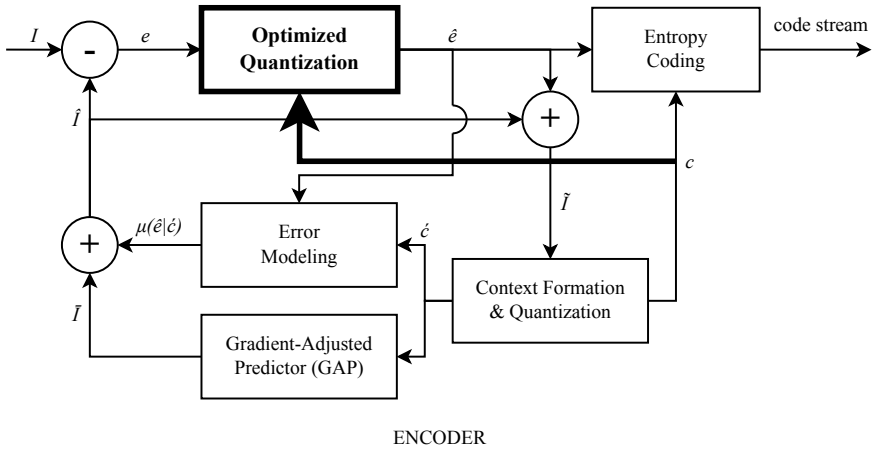


Figure 3.1: Schematic description of proposed method.

weighted directed acyclic graph. Section 3.4 gives an evaluation of the computational complexity of solving the problem. Finally, Section 3.5 explains how the optimal solution is obtained for a wide range of bit rates by varying the Lagrangian multiplier, and the attainable range of bit rates for each ℓ_∞ constraint of τ.

3.1 ℓ_∞-constrained Scalar Quantizer

A quantizer maps the source alphabet into a smaller set of reproduction values. In our case, the source alphabet is a finite set of prediction residues $\mathcal{E} = \{e_n\}_{n=1}^{N}$, where $\{e_1 < e_2 < \cdots < e_N\}$. For raw input images using B bits per pixel, the set of prediction residues $\mathcal{E}$ has $N = 2^{B+1} - 1$ possible integer values ranging from $e_1 = -(2^B - 1)$ to $e_N = 2^B - 1$.

The encoder of a scalar quantizer is described by the partition that segments the source alphabet into a set of non-overlapping contiguous codecells. In other words,

the encoder partition $\mathcal{P}$ can be defined as follows

$$\mathcal{P} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_K\} \quad \text{for some } 1 \leq K \leq N, \tag{3.1}$$

where

$$\begin{aligned} \mathcal{C}_i &= (a_{i-1}, a_i] = \{e_n \mid a_{i-1} < n \leq a_i\} \\ \text{with} \quad & 0 \leq a_{i-1} < a_i \leq N, \quad 1 \leq i \leq K, \end{aligned} \tag{3.2}$$

and $a_0 = 0$ and $a_K = N$.

The decoder of the quantizer, on the other hand, is described by the set of reproduction codewords $\{x_i \mid 1 \leq i \leq K\}$. Every alphabet symbol in codecell $\mathcal{C}_i$ is mapped to the reproduction codeword x_i. In an ℓ_∞-constrained quantizer with a maximum error bound of τ per symbol, the reproduction codeword must satisfy the condition $|e_n - x_i| \leq \tau$, for all $e_n \in \mathcal{C}_i$. On the other hand, keeping in mind the optimization criterion of minimizing the ℓ_2 distortion we determine x_i as follows

$$x_i = \underset{x \in \mathcal{E}, |e_n - x| \leq \tau, e_n \in \mathcal{C}_i}{\arg\min} \sum_{e_n \in \mathcal{C}_i} p(e_n)(e_n - x)^2, \tag{3.3}$$

where $p(e_n)$ is the probability of the prediction residue e_n.

We know that without the ℓ_∞-constraint of $|e_n - x| \leq \tau$ for all $e_n \in \mathcal{C}_i$ in (3.3), the solution to the above minimization problem is simply equal to the centroid of the codecell (Max, 1960; Lloyd, 1982), i.e.,

$$\mu(\mathcal{C}_i) = \sum_{e_n \in \mathcal{C}_i} p(e_n) \frac{e_n}{p(\mathcal{C}_i)}, \tag{3.4}$$

where $p(\mathcal{C}_i) = \sum_{e_n \in \mathcal{C}_i} p(e_n)$ is the probability of codecell $\mathcal{C}_i$.

First of all, for the solution to (3.3) to exist, the size of all codecells must be limited to at most $(2\tau + 1)$. This means that for each codecell $\mathcal{C}_i = (a_{i-1}, a_i]$, the condition $(a_i - a_{i-1}) \leq (2\tau + 1)$ has to be satisfied. Furthermore, to achieve the ℓ_∞-constraint requirement, we must also ensure that the reproduction codeword x_i of codecell $\mathcal{C}_i$ is at distance at most τ from the values at the boundaries of the codecell. In other words, we must have $e_{a_i} - \tau \leq x_i \leq (e_{a_{i-1}} + 1) + \tau$. Notice that when the centroid $\mu(\mathcal{C}_i)$ exceeds distance τ from the boundaries of the codecell, taking the reproduction codeword x_i as the closest point to the centroid within distance τ is the optimal solution to (3.3). This is because the objective function in (3.3) is a quadratic function that is symmetrical around its point of minimum, i.e., around the centroid $\mu(\mathcal{C}_i)$. Therefore, the optimal solution to (3.3) is given by

$$x_i = \begin{cases} (e_{a_{i-1}} + 1) + \tau, & \text{if } (\mu(\mathcal{C}_i) - (e_{a_{i-1}} + 1)) > \tau \\ e_{a_i} - \tau, & \text{if } (e_{a_i} - \mu(\mathcal{C}_i)) > \tau \\ \mu(\mathcal{C}_i), & \text{otherwise.} \end{cases} \tag{3.5}$$

To ease coding and save on memory resources, we round x_i to the nearest integer and limit x_i such that $e_{a_{i-1}} + 1 \leq x_i \leq e_{a_i}$. The algorithmic version of the codeword optimization is also provided in Appendix A.2 for further clarification.

3.2 Optimization Problem Formulation

By optimizing the reproduction codewords for each encoder partition via (3.5), the ℓ_2 distortion and the output entropy corresponding to a quantizer become only functions

of the encoding partition. Let us denote the ℓ_2 distortion and the output entropy for each codecell $\mathcal{C}_i$ as

$$d(\mathcal{C}_i) = \sum_{e_n \in \mathcal{C}_i} p(e_n)(e_n - x_i)^2 \tag{3.6}$$

$$\text{and} \quad r(\mathcal{C}_i) = -p(\mathcal{C}_i) \log_2 p(\mathcal{C}_i) \tag{3.7}$$

respectively. Then the ℓ_2 distortion and the output entropy corresponding to a quantizer with encoder partition $\mathcal{P}$ are

$$D(\mathcal{P}) = \sum_{\mathcal{C} \in \mathcal{P}} d(\mathcal{C}) \tag{3.8}$$

$$\text{and} \quad R(\mathcal{P}) = \sum_{\mathcal{C} \in \mathcal{P}} r(\mathcal{C}) \tag{3.9}$$

Now let us denote by $\mathcal{P}_m$ the encoder partition corresponding to the scalar quantizer for coding context c_m, where $1 \leq m \leq M$ and $M = 8$ for near-lossless CALIC. Subsequently, let D_T and R_T respectively be the expected ℓ_2 distortion and rate of all quantizers over all M contexts as follows

$$D_T = \sum_{m=1}^{M} q(c_m) D(\mathcal{P}_m) \tag{3.10}$$

$$\text{and} \quad R_T = \sum_{m=1}^{M} q(c_m) R(\mathcal{P}_m) \tag{3.11}$$

where $q(c_m)$ is the probability of context c_m. It is important to note that in the computation of $D(\mathcal{P}_m)$ and $R(\mathcal{P}_m)$ using (3.8) and (3.9), respectively, the probability $p(e_n)$ has to be replaced by the conditional probability of residual e_n conditioned on context c_m.

After having established the above notations we can now formulate our task as a constrained optimization problem that minimizes the total ℓ_2 distortion D_T subject to an upper bound ρ on the average rate R_T. In other words,

$$\begin{aligned} \min_{\{\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_M\}} \quad & D_T \\ \text{subject to} \quad & R_T \leq \rho \end{aligned} \tag{3.12}$$

where the optimization is performed over all possible M-tuples of partitions $\mathcal{P}_1$, $\mathcal{P}_2, \ldots, \mathcal{P}_M$ with codecells of maximum size $2\tau + 1$.

As it is common practice in the literature on entropy-constrained quantizer design, we convert the constrained problem (3.12) to the unconstrained problem of minimizing the associated Lagrangian (Muresan and Effros, 2008; Chou *et al.*, 1989). Thus, to achieve an optimal M-tuple of partitions $\{\mathcal{P}_1^*, \mathcal{P}_2^*, \ldots, \mathcal{P}_M^*\}$, we solve

$$\{\mathcal{P}_1^*, \mathcal{P}_2^*, \ldots, \mathcal{P}_M^*\} = \underset{\{\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_M\}}{\arg\min} \{D_T + \gamma R_T\}. \tag{3.13}$$

The solution to problem (3.13) corresponds to a point $({R_T}^*, {D_T}^*)$ on the lower convex hull of the set $\mathcal{RD}$ of all possible planar points of coordinates (R_T, D_T), such that the slope of a tangent to the set $\mathcal{RD}$ drawn through $({R_T}^*, {D_T}^*)$ is equal to $-\gamma$ (Luenberg, 1969).

Substituting (3.10) and (3.11) back into (3.13) and rearranging, it follows that (3.13) is equivalent to

$$\sum_{m=1}^{M} \Big\{ q(c_m) \min_{\mathcal{P}_m} J(\mathcal{P}_m, \gamma) \Big\}, \tag{3.14}$$

where $J(\mathcal{P}_m, \gamma) = D(\mathcal{P}_m) + \gamma R(\mathcal{P}_m)$. This shows that we can minimize $\{D_T + \gamma R_T\}$ in (3.13) by individually minimizing $J(\mathcal{P}_m, \gamma)$ for each context c_m. Note also that since the ℓ_2 distortion $D(\mathcal{P}_m) = \sum_{\mathcal{C}\in\mathcal{P}_m} d(\mathcal{C})$ and rate $R(\mathcal{P}_m) = \sum_{\mathcal{C}\in\mathcal{P}_m} r(\mathcal{C})$ are additive over codecells, the Lagrangian $J(\mathcal{P}_m, \gamma)$ is also additive over codecells, in other words, the following holds

$$J(\mathcal{P}_m, \gamma) = \sum_{\mathcal{C}\in\mathcal{P}_m} j(\mathcal{C}, \gamma), \tag{3.15}$$

$$\text{where} \quad j(\mathcal{C}, \gamma) = d(\mathcal{C}) + \gamma r(\mathcal{C}). \tag{3.16}$$

3.3 Solution Using the Minimum Weight Path Model

Due to the additive nature of the Lagrangian cost shown in (3.15), the task of minimizing (3.15) can simply be viewed as a single-source minimum weight path problem in a weighted directed acyclic graph (WDAG). Each $n \in \{0, 1, \ldots, N\}$ represents a node in the graph, and each pair (x, y) represents an edge that extends from node x to node y. The edge (x, y) symbolizes a possible codecell $\mathcal{C} = (x, y]$, thus only edges (x, y) with $y - x \leq 2\tau + 1$ are allowed. The Lagrangian cost $j(\mathcal{C}, \gamma)$ obtained using (3.16) determines the weight of the edge designating codecell $\mathcal{C}$. We denote by $w(x, y]$ the weight of the edge (x, y), i.e., $w(x, y] = j(\mathcal{C}, \gamma)$, where $\mathcal{C} = (x, y]$. A simple example of a WDAG for constraint $\tau = 1$ is shown in Fig. 3.2. The arrows represent the directed edges for all possible codecells and are labeled with their respective weights. Note that an edge from node 0 to node 4 is not possible because the codecell it represents would exceed the maximum size of $2\tau + 1 = 3$.

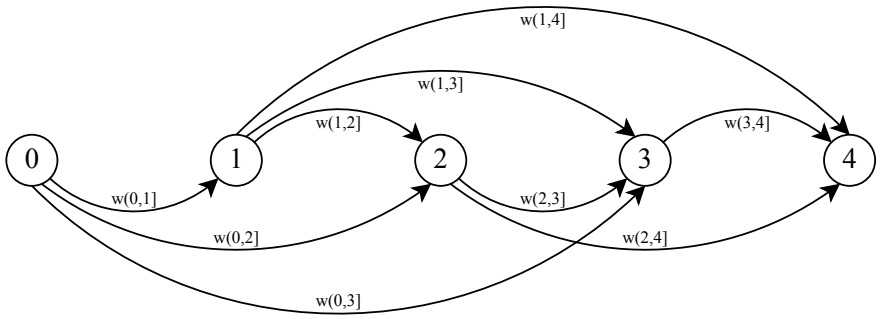


Figure 3.2: All possible edges for $\tau = 1$ in a WDAG with nodes $n \in \{0, 1, 2, 3, 4\}$.

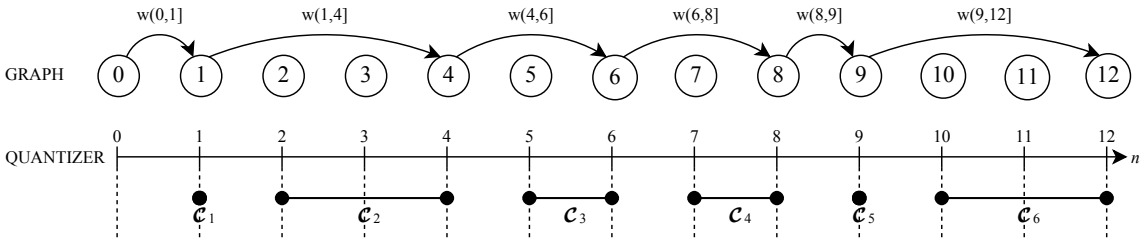


Figure 3.3: An example of a path, made out of a sequence of six edges in a graph, and its corresponding quantizer partition, made out of six contiguous codecells $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_6\}$.

A path in the graph is a sequence of connected edges and the weight of a path is the sum of the weights of all edges which make up that path. It is clear that any path in the graph from 0 to N is in unique correspondence with a partition $\mathcal{P}_m$. Furthermore, from (3.15), we see that the Lagrangian cost $J(\mathcal{P}_m, \gamma)$ of the partition equals the weight of the path. Hence, a path signifies a partition, the weight of a path equals the Lagrangian cost of the partition the path signifies, and minimizing the weight of a path is equivalent to minimizing the Lagrangian cost of the partition in (3.15) that corresponds to that path. Fig. 3.3 illustrates an example of a path in a graph from node 0 to node 12 to partition prediction errors $\{e_n \mid 0 < n \leq 12\}$. In the example,

the path is made out of six edges, representing six codecells $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_6\}$, and the resulting weight of the path is $w(0,1] + w(1,4] + w(4,6] + w(6,8] + w(8,9] + w(9,12]$.

Let $W(x,y]$ be the weight of the minimum weight path from x to y for $0 \leq x < y \leq N$. It is important to observe that if a minimum weight path from 0 to z stops at intermediate node y, then its path from 0 to y must also be a minimum weight path. We can thus find $W(0,z]$, the weight of the minimum weight path from 0 to z, by first computing the minimum weight path from 0 to y. In our case, since all codecells have a size limit of $2\tau+1$ to obey our ℓ_∞ constraint, we need only to compare at most $2\tau+1$ number of weights to obtain $W(0,z]$ for any value of z. Hence, the weight of the minimum weight path from 0 to z is simply

$$W(0,z] = \begin{cases} w(0,z], & \text{if } z = 1 \\ \min\left\{w(0,z], \min_{y>0, z-(2\tau+1)\leq y<z}\left\{W(0,y] + w(y,z]\right\}\right\}, & \text{if } 1 < z \leq (2\tau+1) \\ \min_{z-(2\tau+1)\leq y<z}\left\{W(0,y] + w(y,z]\right\}, & \text{otherwise.} \end{cases} \tag{3.17}$$

Finally, to obtain the minimum weight $W(0,N]$ we compute all minimum weights $W(0,z]$ in increasing order of z, from 1 to N, by using (3.17). A partition $\mathcal{P}_m^*$ with nodes $\{a_i\}_{i=0}^K$ from the minimum weight path of weight $W(0,N]$ would therefore give the optimal quantizer required for every context c_m. To avoid any ambiguity in our explanation, the algorithm is formally described in Appendix A.1.

3.4 Computational Cost

The computation time to obtain the weight of each edge depends linearly on the size of the codecell represented by each edge . With a size limit of $2\tau+1$ on every codecell, the complexity is therefore $O(\tau)$.

Since there are at most $(2\tau+1)$ possible edges ending at N nodes, the weights of at most $N(2\tau+1)$ edges are computed and stored. Subsequently, the minimum weight path to N nodes are found by considering $(2\tau+1)$ weights for every node, except for the first 2τ nodes, where for every n-th node, only n weights are considered.

For all purposes of this thesis however, only Laplacian distributions that are symmetrical around zero are considered. The symmetric distributions give symmetric distortion and entropy measures, and hence, the optimal quantizers achieved should also be symmetric (Sullivan, 1996). To obtain the optimal midtread scalar quantizer with symmetrical codecells with respect to zero, we need only to compute and store the weights of $(2\tau+1)\left(\frac{N+1}{2}-1\right)+(\tau+1)$ edges, and find the minimum weight path to $\left(\frac{N+1}{2}+\tau\right)$ nodes by considering $(2\tau+1)$ path weights for each node, except for the first 2τ nodes and the last $(\tau+1)$ nodes after zero, where fewer number of path weights are considered.

Thus for all M contexts, at most $M(\tau+1)\left[(2\tau+1)(\frac{N+1}{2})+\tau\right]$ loops are executed resulting in $O(\tau^2MN)$ complexity. Since τ and M are upper bounded by a constant much smaller than the value of N, we can further approximate the complexity to be linear, $O(N)$.

3.5 Effects of Varying γ and τ

For each ℓ_∞ error bound of τ, γ in (3.13) controls the tradeoff between the two terms of the Lagrangian cost function, the total ℓ_2 distortion D_T and the total output entropy R_T, which cannot both be minimized simultaneously. A lower value of γ leads to a smaller weight on R_T, and consequently shifts the minimization priority to minimizing D_T. Therefore, a lower value of γ allows us to achieve lower ℓ_2 distortion (or a higher PSNR) at a higher rate, and vice versa.

As observed in our experiments, to obtain the smallest total output entropy R_T for an ℓ_∞ constraint of τ, the optimal quantization solutions $\{\mathcal{P}_1^*, \mathcal{P}_2^*, \ldots, \mathcal{P}_M^*\}$ are very close or even identical to the uniform quantizers in near-lossless CALIC, which have the largest possible step size of $(2\tau + 1)$.

The bigger the value of τ, the larger the maximum step size $(2\tau + 1)$ of the quantizers, the lower the smallest attainable total output entropy $R_0(\tau)$; hence $R_0(\tau + 1) < R_0(\tau)$. This also means that an entropy target of R, for some $R \geq R_0(\tau)$, is achievable by the M-tuple of quantizers optimized with any ℓ_∞ error bound of $\tau^* \geq \tau$. Note that though the quantizers are optimized to obtain the same average entropy R, the different ℓ_∞ constraints might give different M-tuple of optimal quantizers.

Chapter 4

Experimental Results and Remarks

A training set of four 8-bit high resolution continuous-tone images, shown in Fig. 4.1, were used to obtain the probability distributions of prediction errors for every context c_m and every value of $\tau \in \{1, 2, \cdots, 8\}$. We point out that the distributions corresponding to different values of τ are generally different since the contexts are obtained differently for each τ.

Further, each of those distributions was approximated with a Laplacian distribution centered at zero. The approximations were done by choosing the Laplacian probability mass functions (pmfs) with the smallest average difference from the actual pmfs of the training set, i.e., by solving

$$\min_{b>0} \left\{ \frac{1}{N} \sum_{n=1}^{N} \left[p_{lap}(e_n) - p_{tset}(e_n) \right] \right\},$$

$$\text{where} \quad p_{lap}(e_n) = \begin{cases} \frac{1}{T} \left[1 - e^{-\frac{1}{2b}} \right], & \text{if } e_n = 0 \\ \frac{1}{2T} \left[e^{-\frac{|e_n|-0.5}{b}} - e^{-\frac{|e_n|+0.5}{b}} \right], & \text{otherwise} \end{cases}$$

$$T = 1 - e^{-\frac{N}{2b}}. \tag{4.1}$$

(a) 1800 × 1800 pixels

(b) 1701 × 1701 pixels

(c) 1700 × 1700 pixels

(d) 1601 × 1601 pixels

Figure 4.1: Training set images.

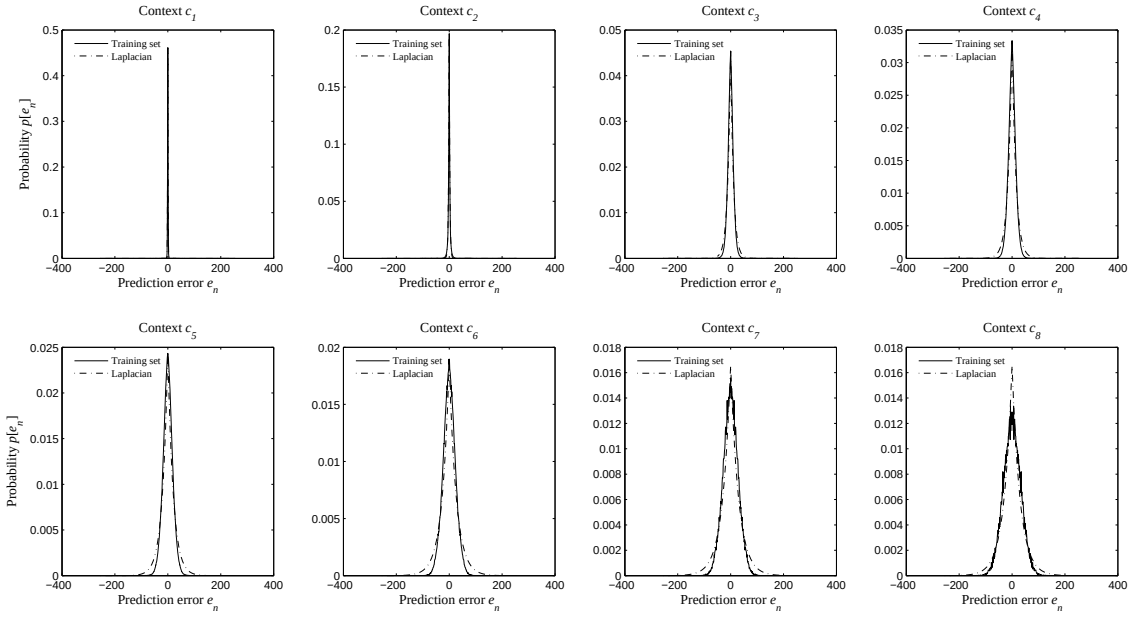


Figure 4.2: Laplacian distributions approximating the distributions of prediction errors from the training set for $\tau = 5$.

$p_{tset}(e_n)$ denotes the actual probability of e_n from the training set. An example of the approximations for $\tau = 5$ are shown in Fig. 4.2, and it can be seen that the Laplacian distributions very closely approximate the distributions of prediction errors. The approximations are necessary to obtain more generalized distributions which take into account random or outlying residue values not found in the training set.

For each $\tau \in \{1, 2, \cdots, 8\}$, we have solved (3.13) for a decreasing sequence of γ values. The pairs (R_T, PSNR_T) corresponding to these solutions are plotted in Fig. 4.3, where $\text{PSNR}_T = 20 \log_{10} \frac{255}{\sqrt{D_T}}$. Notice that, for each τ, rates in the interval $[R_0(\tau), R_0(0)]$ can also be achieved with $\tau' > \tau$, as claimed in Section 3.5 of Chapter 3. Further, in order to proceed to testing the proposed coder on real images, we have selected for each τ, only those M-tuples of codebooks corresponding to entropies R_T satisfying the condition $R_0(\tau) \leq R_T < R_0(\tau - 1)$. Fig. 4.4 plots the resulting PSNR_T

versus R_T of the M-tuples of codebooks selected from Fig. 4.3.

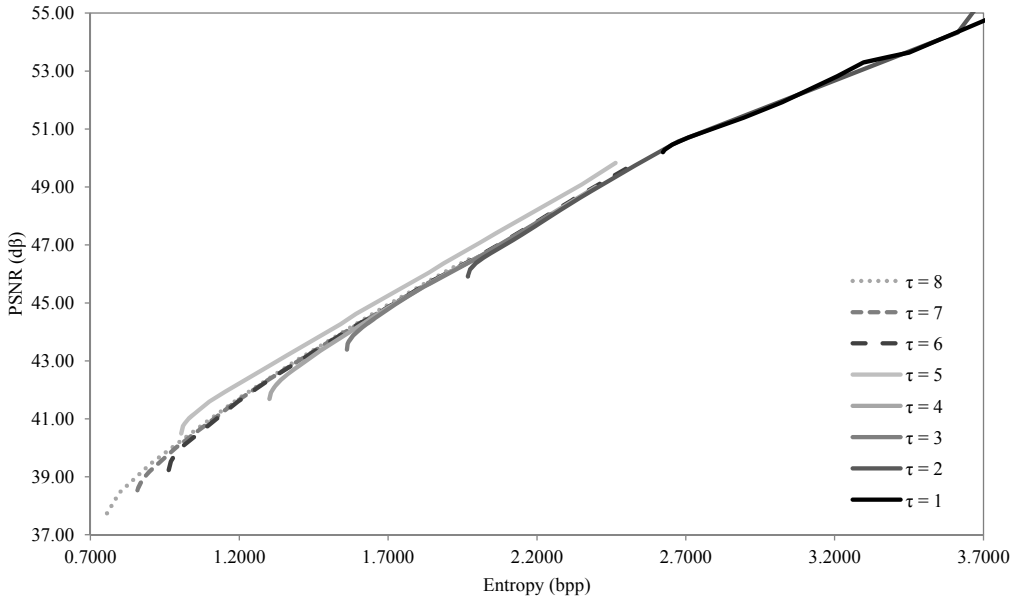


Figure 4.3: PSNR vs. entropy plots of optimal quantizers trained on training set images for different values of τ.

The test images used in our simulations are shown in Fig. 4.5. They were chosen such that a wide range of textures are covered. We point out that the last three images were cropped out of images from the training set.

Fig. 4.6 plots the ℓ_∞ error bound versus output entropy to demonstrate the superiority of the ℓ_∞ error bound of our proposed solution, which is always lower than that of JPEG 2000 for all achievable bit rates. Fig. 4.7 plots the PSNR versus output entropy to further demonstrate the ℓ_2 performance in terms of PSNR as compared to the original near-lossless CALIC and JPEG 2000. The original near-lossless CALIC was only able to achieve a small discrete set of (R, D) points for each value of τ. We have proven that when rate flexibility is allowed for a given ℓ_∞ error bound of τ, the

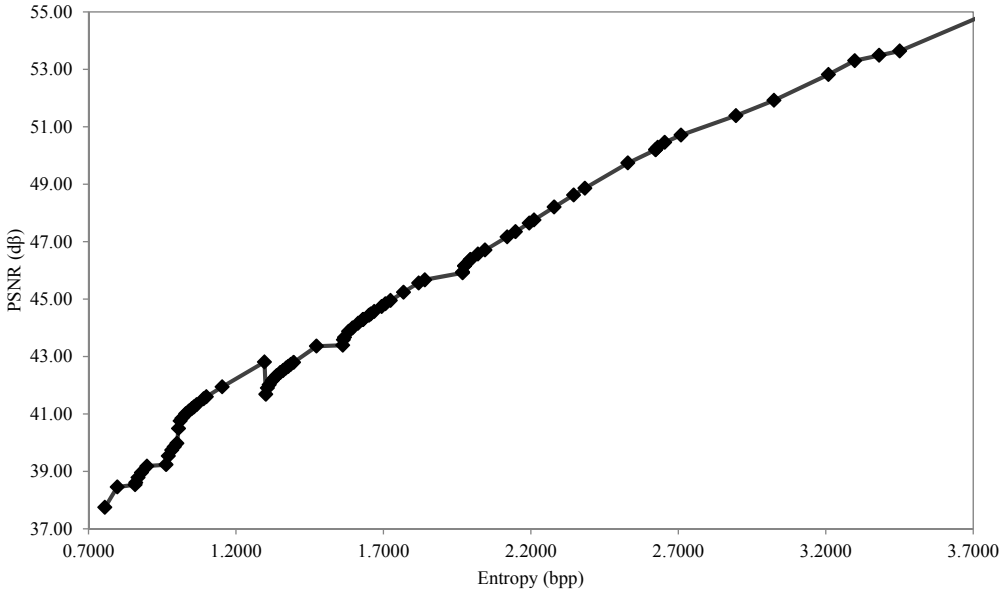


Figure 4.4: PSNR vs. entropy plot of optimal quantizers chosen from Fig. 4.3.

ℓ_2 performance of our proposed solution surpasses that of near-lossless CALIC. The resulting plot also demonstrates that our proposed solution outperforms JPEG 2000 above certain threshold bit rates, which can be as low as 1.1 bpp.

The authors of (Wu and Bao, 2000) experimented on incorporating the DPCM trellis implemented in (Ke and Marcellin, 1998) into near-lossless CALIC, and concluded that it did not offer appreciable compression gains despite the high computational complexity incurred. Since our work is able to achieve the compression performances of near-lossless CALIC and better with many more bit rates, we will not further attempt to compare our work to the work in (Ke and Marcellin, 1998).

(a) Hair (256×256 pixels)

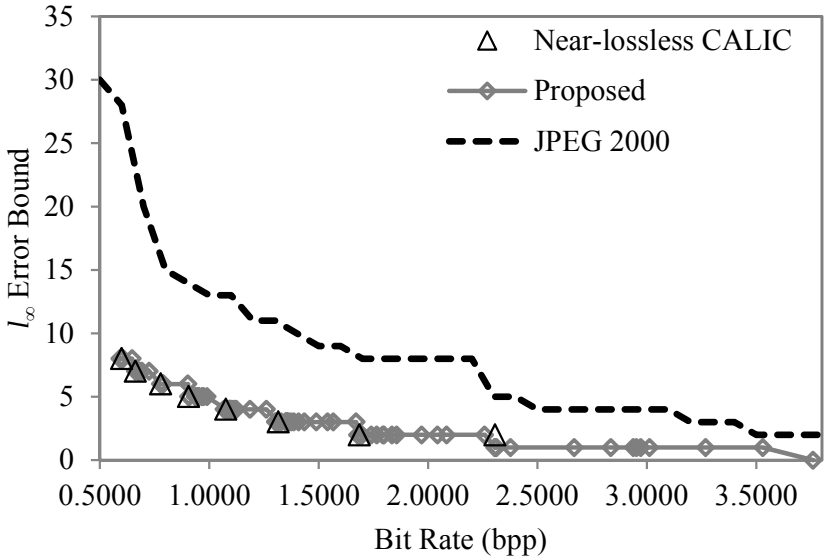
(b) Plant (400×600 pixels)

(c) Flowers (256×256 pixels)

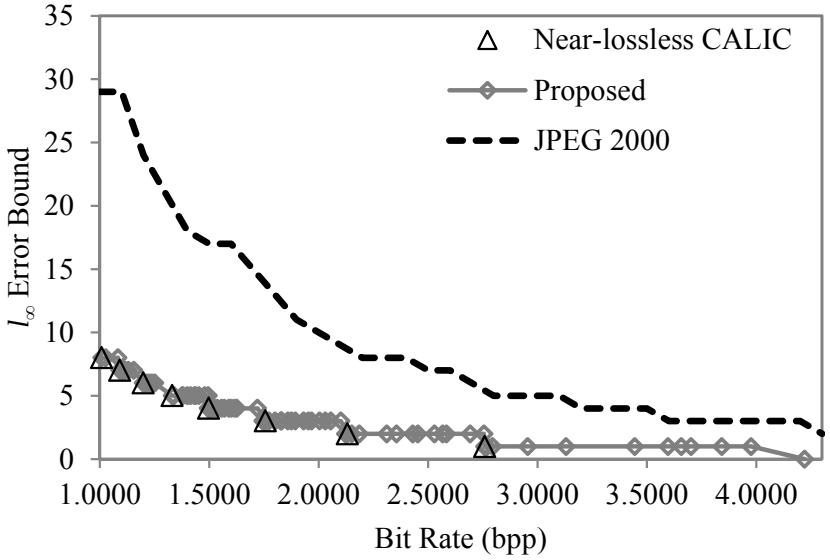
(d) Plants (256×256 pixels)

(e) Fruits (600×400 pixels)

Figure 4.5: Test images.

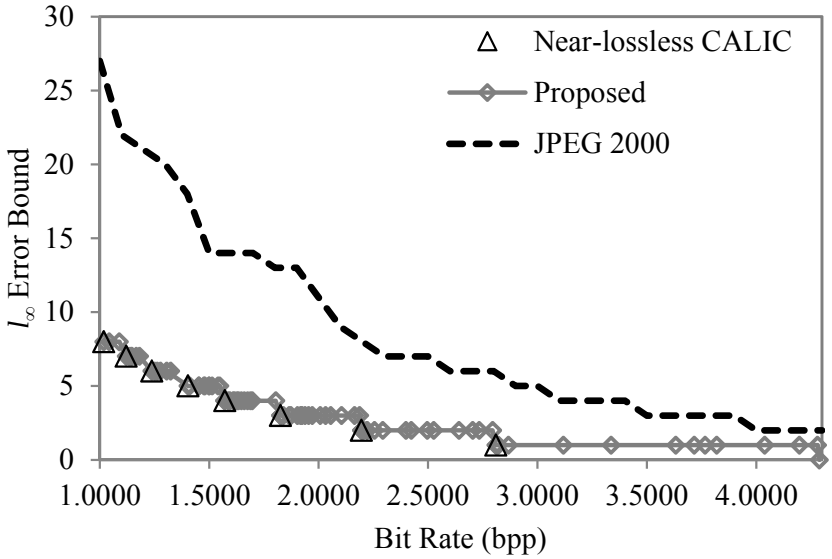


(a) Hair

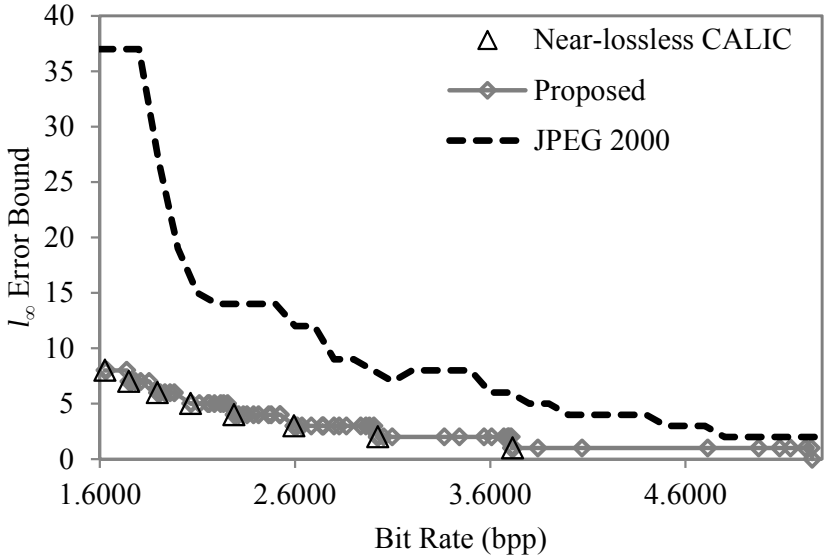


(b) Plant

Figure 4.6: ℓ_∞ error bound of images (a) Hair and (b) Plant compressed at different rates.

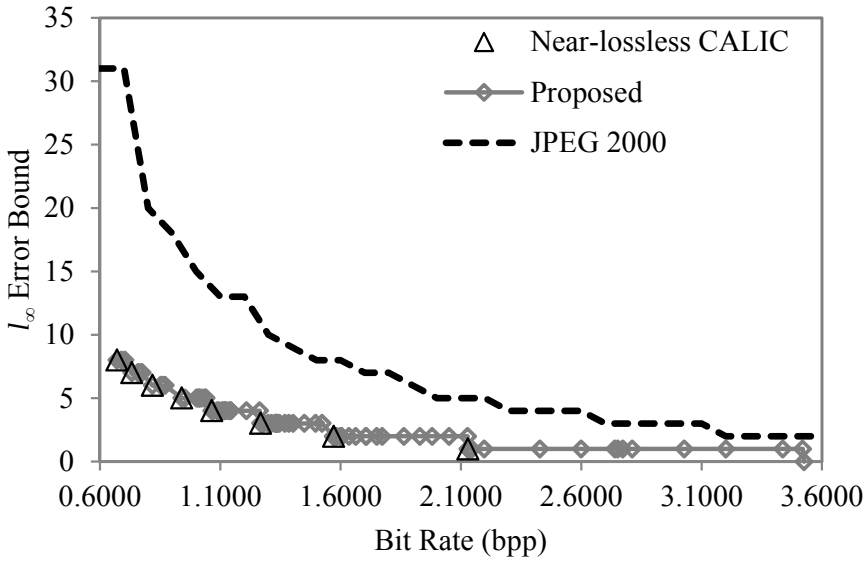


(c) Flowers



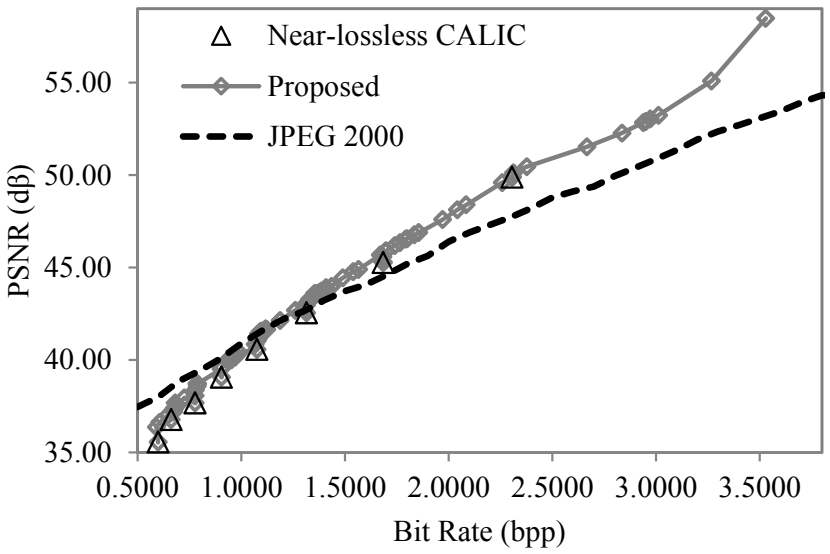
(d) Plants

Figure 4.6: ℓ_∞ error bound of images (c) Flowers and (d) Plants compressed at different rates.



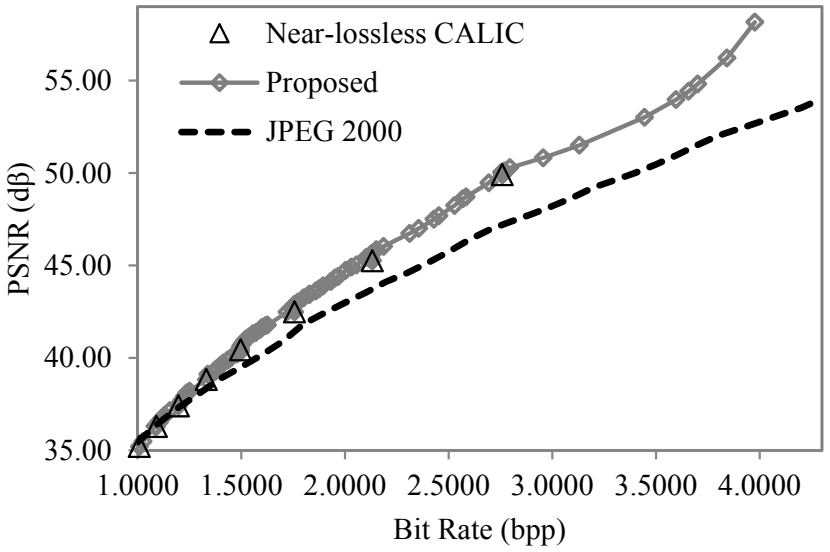
(e) Fruits

Figure 4.6: ℓ_∞ error bound of image (e) Fruits compressed at different rates.

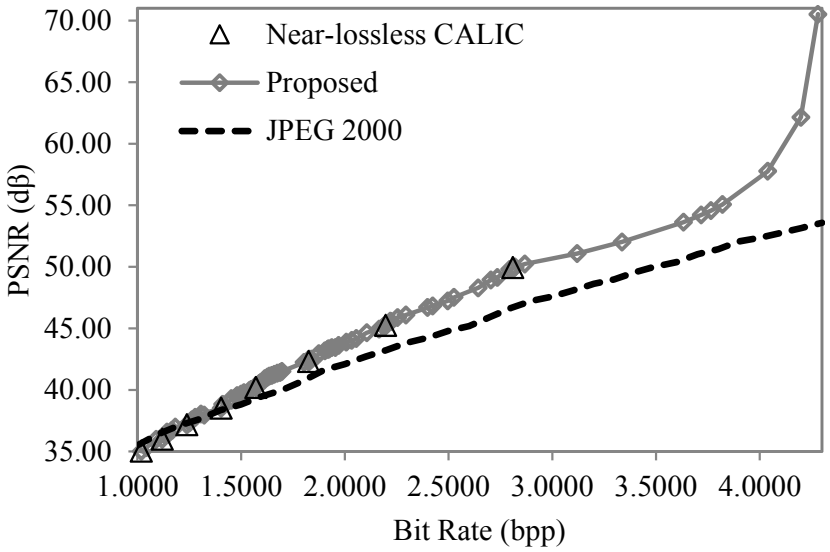


(a) Hair

Figure 4.7: PSNR of image (a) Hair compressed at different rates.

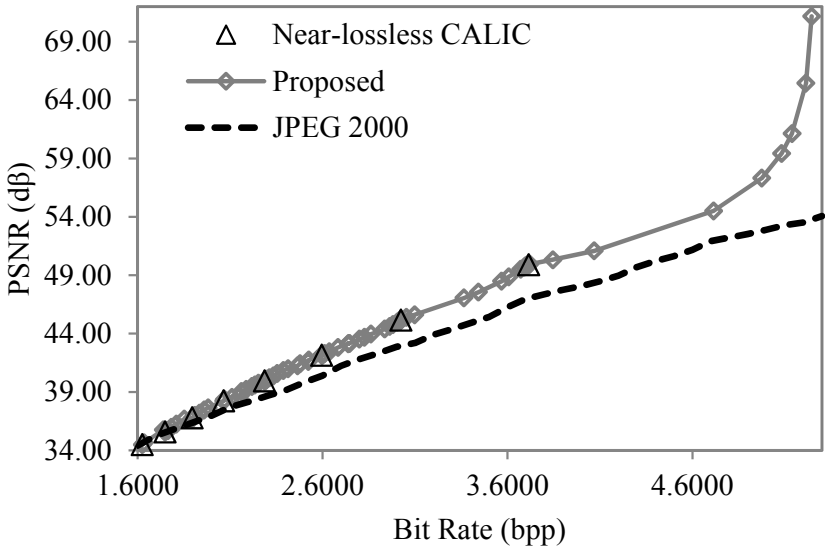


(b) Plant

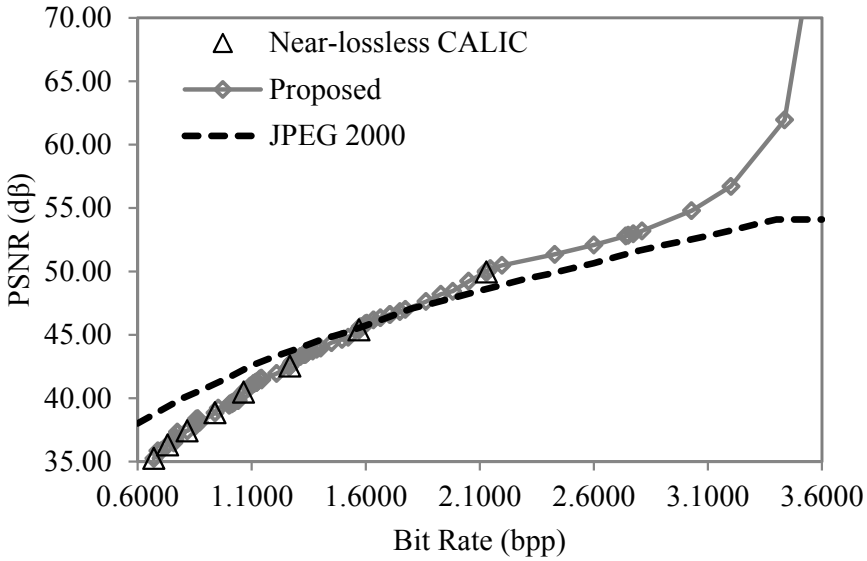


(c) Flowers

Figure 4.7: PSNR of images (b) Plant and (c) Flowers compressed at different rates.



(d) Plants



(e) Fruits

Figure 4.7: PSNR of images (d) Plants and (e) Fruits compressed at different rates.

4.1 Probability Distribution for Optimization

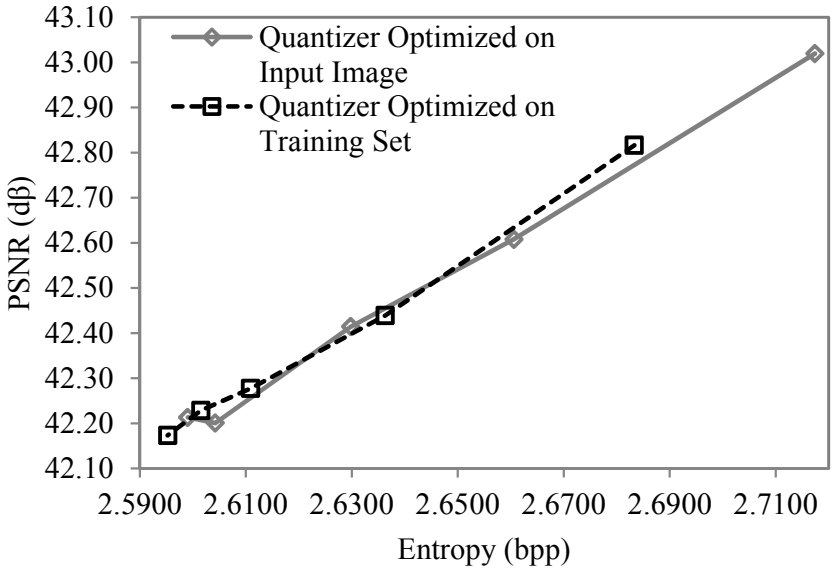
A difficulty in the optimization lies in the problem of finding the best possible distribution to optimize to. While the training set may be chosen to be generic enough to represent all possible images, the distribution of each input image will still vary to some degree from the distribution of the training set. As such, we cannot guarantee optimality for each input image as much as we can for the training set used for optimization.

We have thus experimented on adaptively optimizing the scalar quantizers based on the probability distributions of the input image, which are approximated by Laplacian distributions. However, the results, as displayed in Fig. 4.8, showed little to no improvement over using scalar quantizers optimized on a training set.

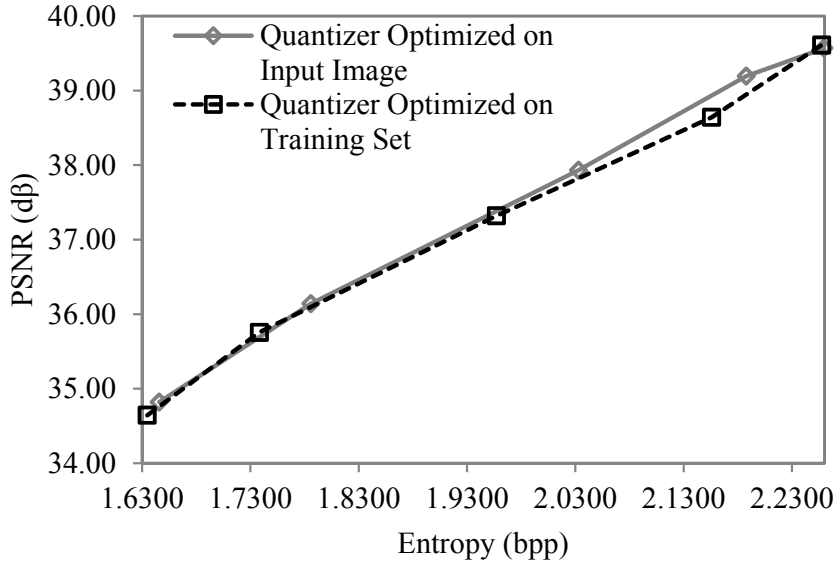
Another main statistical mismatch in the probability distribution for optimization is due to the nature of the predictive coding in near-lossless CALIC. Each prediction error e is computed based on a knowledge of reconstructed neighbouring pixels as shown below.

$$e = I - \left\{ \mathrm{GAP}(\tilde{\mathbf{I}}_{neighbour}) + \mu(\hat{e}|\acute{c}) \right\}, \tag{4.2}$$

where $\tilde{\mathbf{I}}_{neighbour}$ is the set of reconstructed neighbouring pixels $\{\tilde{I}_n, \tilde{I}_{ne}, \tilde{I}_{nn}, \tilde{I}_{nne}, \tilde{I}_{nw}, \tilde{I}_w, \tilde{I}_{ww}\}$ with positions shown in Fig. 2.2, GAP is the gradient-adjusted prediction operator, and $\mu(\hat{e}|\acute{c})$ is the conditional sample mean of the quantized prediction errors; all previously described in Chapter 2. Each neighbouring pixel is reconstructed from their respective predicted value and quantized prediction error, i.e., $\tilde{I}_{neighbour} = \hat{I}_{neighbour} + \hat{e}_{neighbour}$, where $\tilde{I}_{neighbour} \in \tilde{\mathbf{I}}_{neighbour}$. Therefore, the distribution of prediction errors e changes depending on the quantizers in the algorithm.



(a) $\tau = 3$



(b) $\tau = 8$

Figure 4.8: Performance comparison for test image 4.5d coded with quantizers optimized with distributions from the input image and distributions from the training set for $\tau = 3$ and $\tau = 8$.

We have implemented the most straightforward approach in obtaining the probability distribution of prediction errors, known as the Open-Loop (OL) approach, that ignores the dependence of the distribution on the quantizers, and simply generates the distribution of prediction errors e using only the original pixel values I of the training set.

$$e = I - \text{GAP}(\mathbf{I}_{neighbour}), \tag{4.3}$$

where $\mathbf{I}_{neighbour}$ is the set of original neighbouring pixels.

In hopes of finding a better representation of the actual distribution for optimization, we proceeded to test other approaches, the Closed-Loop (CL) approach proposed in (Cuperman and Gersho, 1985) and the Asymptotic Closed-Loop (ACL) approach proposed in (Khalil *et al.*, 2001). The CL approach starts off with the OL approach for the first iteration $i = 1$ to obtain the first set of M optimal quantizers, which we will denote as $\mathbf{Q}^{(1)}$. Then, with those quantizers, further iterations $i = 2, 3, \ldots$ are preformed with the CL approach as shown below in (4.4) to obtain a hopefully more accurate distribution of prediction errors e.

$$\begin{aligned}
& e^{(i)} = I - \text{GAP}(\tilde{\mathbf{I}}^{(i)}_{neighbour}), \\
\text{where} \quad & \tilde{\mathbf{I}}^{(i)}_{neighbour} = \{\tilde{I}^{(i)}_{n}, \tilde{I}^{(i)}_{ne}, \tilde{I}^{(i)}_{nn}, \tilde{I}^{(i)}_{nne}, \tilde{I}^{(i)}_{nw}, \tilde{I}^{(i)}_{w}, \tilde{I}^{(i)}_{ww}\} \\
& \tilde{I}^{(i)} = \text{GAP}(\tilde{\mathbf{I}}^{(i)}_{neighbour}) + Q^{(i-1)}\left(I - \text{GAP}(\tilde{\mathbf{I}}^{(i)}_{neighbour})\right) \\
& Q^{(i-1)} \in \mathbf{Q}^{(i-1)} \\
& i = 2, 3, \ldots
\end{aligned} \tag{4.4}$$

For each iteration $i > 1$, $\mathbf{Q}^{(i-1)}$ is the set of optimal quantizers found for distribution $e^{(i-1)}$. The ACL approach, on the other hand, only generates prediction errors $e^{(i)}$

using predictions from the previous iteration $(i-1)$ as shown in Fig. 4.9 below and (4.5).

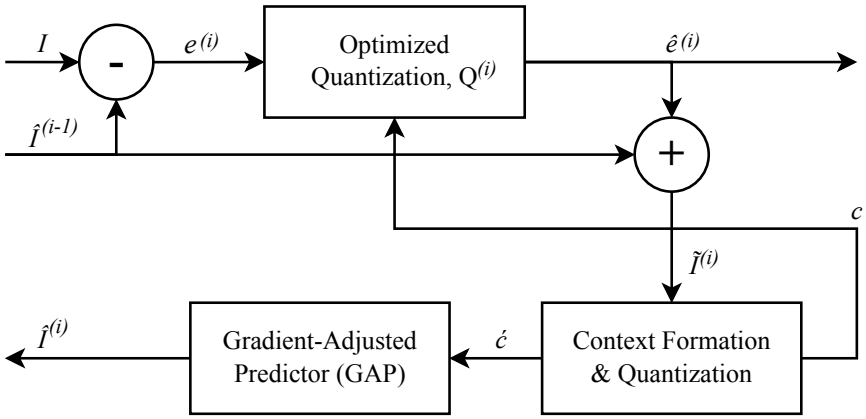


Figure 4.9: Schematic description of how the distribution of prediction errors $e^{(i)}$ is obtained for each iteration i of the ACL approach, where $\hat{I}^{(i)} = \text{GAP}(\tilde{\mathbf{I}}^{(i)}_{neighbour})$.

$$
\begin{aligned}
& e^{(i)} = I - \text{GAP}(\tilde{\mathbf{I}}^{(i-1)}_{neighbour}), \\
\text{where} \quad & \tilde{\mathbf{I}}^{(i-1)}_{neighbour} = \{\tilde{I}^{(i-1)}_{n}, \tilde{I}^{(i-1)}_{ne}, \tilde{I}^{(i-1)}_{nn}, \tilde{I}^{(i-1)}_{nne}, \tilde{I}^{(i-1)}_{nw}, \tilde{I}^{(i-1)}_{w}, \tilde{I}^{(i-1)}_{ww}\} \\
& \tilde{I}^{(i)} = \text{GAP}(\tilde{\mathbf{I}}^{(i-1)}_{neighbour}) + Q^{(i)}\left(I - \text{GAP}(\tilde{\mathbf{I}}^{(i-1)}_{neighbour})\right) \\
& Q^{(i)} \in \mathbf{Q}^{(i)} \\
& i = 1, 2, \ldots
\end{aligned} \tag{4.5}
$$

The approach starts off with $\tilde{\mathbf{I}}^{(0)}_{neighbour} = \{I_n, I_{ne}, I_{nn}, I_{nne}, I_{nw}, I_w, I_{ww}\}$, which are the original pixel values I from the training set, and $\mathbf{Q}^{(1)}$ as the set of quantizers optimized for the distribution from the OL approach. For subsequent iterations $i > 1$, $\mathbf{Q}^{(i)}$ is the set of quantizers optimized for distribution $e^{(i)}$.

Table 4.1 displays the performance comparisons between the OL, CL and ACL approaches for test image 4.5e with $\tau = 6$ and $\gamma = 20.2$. It can be seen that the OL

approach still gives the best distribution of prediction errors for optimization.

Table 4.1: Performance comparison for test image 4.5e coded with quantizers optimized based on distributions from the Open-Loop (OL), Closed-Loop (CL) and Asymptotic Closed-Loop (ACL) approaches for $\tau = 6$ and $\gamma = 20.2$.

	Iteration i	PSNR(dB)	Rate(bpp)	ℓ_2 distortion	Weight
OL	1	38.40	0.86	9.40	26.8478
CL	2	38.29	0.87	9.64	27.1899
	3	38.15	0.87	9.96	27.5892
ACL	1	38.18	0.86	9.89	27.3448
	2	38.05	0.87	10.18	27.8477
	3	38.15	0.87	9.96	27.5892

4.2 Context Quantization Optimization

As previously mentioned in Chapter 2, the scalar quantization performed on error energy estimator Δ to form coding context c was optimized to minimize the conditional entropy of prediction errors. Rather than just minimizing the entropy, we also considered a scheme to optimize the context quantization such that the Lagrangian, i.e. the weighted sum of both the ℓ_2 distortion and the entropy, is minimized.

We optimize the context quantizer with a WDAG method similar to that described in Section 3.3 with the exception of the ℓ_∞ constraint, and the addition of a limit to the number of codecells to only M to obtain M coding contexts c. The nodes of the graph now represent all possible values of Δ to be quantized, where $0 \leq \Delta \leq N$. Without the ℓ_∞ constraint, there is no longer a size limit to each codecell. For a progressively increasing number of codecells c, from 1 to M, the weights of the smallest weighted paths $W_c(0, z]$ are found for all endpoint nodes z, from c to $N+1-(M-c)$. In other

words,

FOR all c from 1 to M

FOR each endpoint z from c to $N+1-(M-c)$

IF $c=1$

$W_c(0,z] = w(0,z];$

ELSE

$W_c(0,z] = \min_{c-1 \le y < z} \{W_{c-1}(0,y] + w(y,z]\};$

ENDIF

ENDFOR

ENDFOR (4.6)

where $M = 8$ and $w(a,b]$ is the Lagrangian weighted sum $J(\mathcal{P},\gamma)$ of the scalar quantizer optimized for all prediction errors e with conditional probabilities $\{p(e|\Delta) \mid a < \Delta \le b\}$.

As discussed in Section 3.4, the weight of each edge $w(a,b]$ can be computed in linear time $O(N)$, where $N = 2^{B+1} - 1$. To optimize the context quantization however, we need to compute N^2 weights, where $N = 7(2^B - 1) + 1$ as the largest possible value of Δ calculated from (2.3). The computation of the weights of all the edges thus requires $O(N^3)$ time. In addition to that, the inner loop in (4.6) that computes the minimum weight paths runs $M(N - M + 2)$ times, out of which, $(M-1)(N-M+2)$ times are required to compare $(N-M+2)$ path weights when $c > 1$. With M as a constant much smaller than N, the computation time to obtain the minimum weight path is therefore $O(N^2)$. This then yields a very

high total complexity of $O(N^3 + N^2)$, or simply $O(N^3)$ for the entire new context quantizer optimization algorithm. Nevertheless, the huge additional complexity did not offer an improved performance in PSNR or bit rate for most images as compared to the original simpler context quantization in near-lossless CALIC. Tables 4.2 and 4.3 display the performance comparisons at $\tau = 3$ for $\gamma = 1.5$ and $\gamma = 2.0$ respectively.

Table 4.2: Performance comparison between the original and new context quantization for $\tau = 3$ and $\gamma = 1.5$, where the weights in **bold** indicate better overall performance.

Test Image	Original Context Quantization			New Context Quantization		
	ℓ_2 Distortion	Rate	Weight	ℓ_2 Distortion	Rate	Weight
Hair	1.3247	1.8996	**4.1741**	1.4128	1.8751	4.2255
Plant	1.1573	2.4617	**4.8499**	1.1956	2.4471	4.8663
Flowers	0.9912	2.6197	**4.9208**	1.0428	2.6048	4.9501
Plants	0.7501	3.6666	6.2501	0.7438	3.6646	**6.2408**
Fruits	1.3244	1.8114	**4.0415**	1.4135	1.7910	4.1000

Table 4.3: Performance comparison between the original and new context quantization for $\tau = 3$ and $\gamma = 2.0$, where the weights in **bold** indicate better overall performance.

Test Image	Original Context Quantization			New Context Quantization		
	ℓ_2 Distortion	Rate	Weight	ℓ_2 Distortion	Rate	Weight
Hair	1.7576	1.6712	**5.1001**	1.8485	1.6421	5.1326
Plant	1.8554	2.1020	**6.0594**	1.8753	2.0968	6.0690
Flowers	1.6970	2.2597	**6.2163**	1.7605	2.2532	6.2669
Plants	1.8460	3.0825	8.0110	1.8112	3.0905	**7.9922**
Fruits	1.7696	1.5938	**4.9572**	1.8636	1.5803	5.0241

4.3 Practical Considerations

The proposed approach implies optimization of the quantizers, but performing it online increases the complexity of the encoder. Therefore, in order to keep the encoding

complexity low, one option is to perform the optimization offline on a training set and store a number of such optimized M-tuples of quantizers at the encoder. The number of values of τ and rates covered can be chosen taking in consideration particular requirements of the application. For each M-tuple of quantizers, the entropy R_T achieved on the training set or the average bit rate obtained on coded images can be additionally stored in order to help estimate the achievable rate for a particular image. If there are some K_m number of codecells for each context c_m, we only need to store the eight values of $\{(K_m - 1)/2\}_{m=1}^{8}$, and half the number of codewords and codecell upper boundaries, $\sum_{m=1}^{8}(K_m - 1)/2$ each, due to the symmetry of the quantizer. For an 8-bit image, the constant values stored in memory thus occupy $(8 + \sum_{m=1}^{8}(K_m - 1))$ bytes in total, where $K_m \leq N$. So an 8-bit image would only require at most a 4kB look-up table for each allowable rate.

Out of the stored M-tuples of codebooks the user chooses one according to the specifications of the applications. This raises the question of how to accurately and effectively estimate the bit rate achieved for a particular image from the value R_T. Unfortunately, we do not have yet a low complexity solution to this problem, which is left for future work. On the other hand, if the specification on the target bit rate is rather a looser qualitative requirement, such as "low", "medium" or "high" bit rate, then selecting an appropriate M-tuple of quantizers could easily be done. For instance, if the specified ℓ_∞ constraint is τ and the used bit rate has to be low, the obvious choice is the M-tuple operating at the smallest rate, i.e., with $R_T = R_0(\tau)$ or close to this rate. On the other hand, if we can afford a high bit rate, the M-tuple operating at rate $R_T \approx R_0(\tau - 1)$ can be chosen, while for moderate bit rate, a value $R_T \approx (R_0(\tau) + R_0(\tau - 1))/2$ can be selected.

In some applications, such as image archiving, however, the encoder can afford high complexity while decoding complexity still has to be low. In such cases, the proposed technique can be used with an accurate rate control by trying several M-tuples of quantizers among the stored ones in a bisection search fashion until a rate close enough to the target rate is achieved. The online optimization with a training set collected from the image at hand can also be incorporated at the encoder.

Chapter 5

Conclusion

To improve upon the higher bit rate performances of both ℓ_2 and ℓ_∞-based image coding techniques already existing in literature, we proposed a new ℓ_2-optimized variant of near-lossless CALIC, which is a benchmark of good image codecs with an ℓ_∞ bound. Our proposed codec allows for a wide range of achievable rates for each ℓ_∞ bound, unlike pure ℓ_∞-based image coding methods with only one achievable rate per ℓ_∞ bound. At the cost of only slight increases in bit rate, the ℓ_2 optimization gives our codec the ability to significantly reduce structured errors, such as speckles and contours commonly found in pure ℓ_∞-based decoded images. Not only are structured artifacts smoothed out by the ℓ_2 optimization, the ℓ_∞ bound in our codec also maintains a maximum error bound on each pixel that is always lower than that of pure ℓ_2-based coded images at the same bit rate.

Near-lossless CALIC employs context-based predictive coding followed by uniform scalar quantization of residual errors. The idea of our proposed approach is to replace the uniform quantizer by context-based ℓ_2-optimized quantizers. The optimization criterion is to minimize the ℓ_2 distortion subject to a constraint on the entropy while

maintaining a strict ℓ_∞ error bound. The proposed approach greatly increases the rate granularity while guaranteeing both high ℓ_2 and minmax fidelity for the whole range of achievable bit rates at no additional computational cost. Experimental results demonstrate that the technique noticeably outperforms JPEG 2000 in minmax fidelity for all achievable bit rates. Furthermore, above certain bit rates, even bit rates as low as 1.1 bpp, the code also outperforms JPEG 2000 in ℓ_2 fidelity without compromising its superior ℓ_∞ error bound.

While our proposed coding method can achieve a wide range of bit rates, it does not guarantee the same bit rate for every image. For each set of ℓ_2-optimized quantizers, the ℓ_∞ error bound is fixed, but the attainable bit rate varies depending on the image being coded. Images with many smooth regions tend to achieve lower rates than images with many sharp contrasts and gradients. Even though the sets of quantizers can easily be ordered based on attainable bit rates, the complexity in searching for the best set of quantizers to obtain the rate closest to a target rate is high. A better solution to accurately and effectively estimate or set the bit rate of each coded image is thus left for future research.

Appendix A

Context-based Quantizer Optimization Algorithm

A.1 Quantizer Optimization

FOR each context c_m

{Calculate Lagrangian $j(\mathcal{C}, \gamma) = w(a, b]$ for each codecell $\mathcal{C}$}

INIT array of codewords $x(\mathcal{C})$, distortions $d(\mathcal{C})$ and entropies $r(\mathcal{C})$

for all possible codecells $\mathcal{C}$ to 0

READ array of $p(e_n|c_m)$ for all $e_n \in \mathcal{E}$

FOR each cell size s from 0 to 2τ

FOR each lower boundary a from 0 to $N - 1 - s$

$b = a + 1 + s$;

$p(\mathcal{C}) = \sum_{e_n \in \mathcal{C}} p(e_n|c_m)$;

IF $p(\mathcal{C}) \neq 0$

$x(\mathcal{C}) = \text{CODEWORD}(a, b, p(\mathcal{C}), p(e_n|c_m), \tau)$;

$d(\mathcal{C}) = \sum_{e_n \in \mathcal{C}} p(e_n|c_m)(e_n - x(\mathcal{C}))^2$;

$r(\mathcal{C}) = -p(\mathcal{C}) \log_2 p(\mathcal{C})$;

ELSE

$x(\mathcal{C}) = \lfloor (b + a + 1)/2 \rfloor$;

$d(\mathcal{C}) = 0$;

$r(\mathcal{C}) = 0$;

ENDIF

$w(a, b] = d(\mathcal{C}) + \gamma r(\mathcal{C})$;

ENDFOR

ENDFOR

{Find minimum weight paths from $-(2^B - 1)$ to e_n}

INIT array of N minimal path weights $W(0, b]$ and lower boundaries a to 0

FOR each z from 1 to N

IF $z > 1$

$W(0, z] = \min_{y>0,\ z-(2\tau+1) \leq y < z} \{W(0, y] + w(y, z]\}$;

$a(z) = y$; {where y is the value achieving the above minimum}

IF $z \leq (2\tau + 1)$ AND $w(0, z] < W(0, z]$

$W(0,z] = w(0,z];$

$a(z) = 0;$

ENDIF

ELSE

$W(0,z] = w(0,z];$

$a(z) = 0;$

ENDIF

ENDFOR

{Find optimal partitions}

INIT number of codecells $K_m = 0$

INIT upper boundary $b = N$

WHILE upper boundary $b > 0$

STORE upper boundary b

STORE codeword x

$K_m = K_m + 1;$

$b = a(b);$

ENDWHILE

STORE number of codecells K_m

ENDFOR (A.1)

A.2 Codeword Optimization

FUNCTION $x = \text{CODEWORD}\ (a, b, p(\mathcal{C}), p(e_n|c_m), \tau)$

$x = \sum_{n=a+1}^{b} p(e_n|c_m)\frac{e_n}{p(\mathcal{C})}$;

$x = \text{round}(x)$;

IF $(x - e_{a+1}) > \tau$

$x = e_{a+1} + \tau$;

ELSEIF $(e_b - x) > \tau$

$x = e_b - \tau$;

ENDIF

IF $x < e_{a+1}$

$x = e_{a+1}$;

ELSEIF $x > e_b$

$x = e_b$;

ENDIF

ENDFUNCTION (A.2)

Bibliography

Bruce, J. D. (1964). *Optimum quantization*. Ph.D. thesis, M. I. T.

Chen, K. and Ramabadran, T. (1994). Near-lossless compression of medical images through entropy-coded dpcm. *Medical Imaging, IEEE Transactions on*, **13**(3), 538 – 548.

Chou, P., Lookabaugh, T., and Gray, R. (1989). Entropy-constrained vector quantization. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, **37**(1), 31 – 42.

Cuperman, V. and Gersho, A. (1985). Vector predictive coding of speech at 16 kbits/s. *Communications, IEEE Transactions on*, **33**(7), 685 – 696.

Dumitrescu, S. and Wu, X. (2004). Algorithms for optimal multi-resolution quantization. *Journal of Algorithms*, **50**(1), 1 – 22.

Dumitrescu, S. and Wu, X. (2005). Optimal two-description scalar quantizer design. *Algorithmica*, **41**, 269 – 287.

Dumitrescu, S. and Wu, X. (2007). Lagrangian optimization of two-description scalar quantizers. *Information Theory, IEEE Transactions on*, **53**(11), 3990 – 4012.

Ke, L. and Marcellin, M. (1998). Near-lossless image compression: minimum-entropy, constrained-error dpcm. *Image Processing, IEEE Transactions on*, **7**(2), 225 – 228.

Khalil, H. and Rose, K. (2003). Predictive vector quantizer design using deterministic annealing. *Signal Processing, IEEE Transactions on*, **51**(1), 244 – 254.

Khalil, H., Rose, K., and Regunathan, S. (2001). The asymptotic closed-loop approach to predictive vector quantizer design with application in video coding. *Image Processing, IEEE Transactions on*, **10**(1), 15 –23.

Lloyd, S. (1982). Least squares quantization in pcm. *Information Theory, IEEE Transactions on*, **28**(2), 129 – 137.

Luenberg, D. G. (1969). *Optimization by Vector Space Methods*. John Wiley & Sons, New York.

Max, J. (1960). Quantizing for minimum distortion. *Information Theory, IRE Transactions on*, **6**(1), 7 – 12.

Muresan, D. and Effros, M. (2008). Quantization as histogram segmentation: Optimal scalar quantizer design in network systems. *Information Theory, IEEE Transactions on*, **54**(1), 344 – 366.

Netravali, A. (1977). On quantizers for dpcm coding of picture signals. *Information Theory, IEEE Transactions on*, **23**(3), 360 – 370.

Said, A. and Pearlman, W. (1996). An image multiresolution representation for lossless and lossy compression. *Image Processing, IEEE Transactions on*, **5**(9), 1303 –1310.

Sharma, D. (1978). Design of absolutely optimal quantizers for a wide class of distortion measures. *Information Theory, IEEE Transactions on*, **24**(6), 693 – 702.

Sullivan, G. (1996). Efficient scalar quantization of exponential and laplacian random variables. *Information Theory, IEEE Transactions on*, **42**(5), 1365 –1374.

Wu, X. (1991). Optimal quantization by matrix searching. *Journal of Algorithms*, **12**(4), 663 – 673.

Wu, X. and Bao, P. (2000). L infin; constrained high-fidelity image compression via adaptive context modeling. *Image Processing, IEEE Transactions on*, **9**(4), 536 – 542.

Wu, X. and Memon, N. (1997). Context-based, adaptive, lossless image coding. *Communications, IEEE Transactions on*, **45**(4), 437 – 444.

Wu, X. and Zhang, K. (1993). Quantizer monotonicities and globally optimal scalar quantizer design. *Information Theory, IEEE Transactions on*, **39**(3), 1049 –1053.

Wu, X., Memon, N., and Sayood, K. (1995). A context-based, adaptive, lossless / nearly-lossless coding scheme for continuous-tone images. *ISO/IEC JTC*.

Wu, X., Zhou, J., and Wang, H. (2011). High-fidelity image compression for high-throughput and energy-efficient cameras. In *Data Compression Conference (DCC), 2011*, pages 433 – 442.

Zandi, A., Allen, J., Schwartz, E., and Boliek, M. (1995). Crew: Compression with reversible embedded wavelets. In *Data Compression Conference, 1995. DCC '95. Proceedings*, pages 212 – 221.