Rate Bounds on SSIM Index of Quantized Image DCT Coefficients

Sumohana S. Channappayya, Alan C. Bovik, Robert W. Heath Jr. and
Constantine Caramanis

Dept. of Elec. & Comp. Engg.,The University of Texas at Austin, Austin, TX 78712-0240, USA

{sumohana, bovik, rheath, cmcaram}@ece.utexas.edu

January 8, 2008

Abstract

In this paper, we derive bounds on the Structural SIMilarity (SSIM) Index as a function of quantization rate for fixed-rate uniform quantization of image discrete cosine transform (DCT) coefficients under the high rate assumption. The space domain SSIM Index is first expressed in terms of the DCT coefficients of the space domain vectors. The transform domain SSIM Index is then used to derive bounds on the average SSIM Index as a function of quantization rate for Gaussian and Laplacian sources. As an illustrative example, uniform quantization of the DCT coefficients of natural images is considered. We show that the SSIM Index between the reference and quantized images fall within the bounds for a large set of natural images. Further, we show using a simple example that the proposed bounds could be very useful for rate allocation problems in practical image and video coding applications.

1 Introduction

The mean squared error (MSE) is a popular distortion measure used in the optimization of a wide variety of image processing algorithms ranging from quantization to restoration. The amenability of the MSE to analysis combined with a lack of competitive perceptual distortion measures contribute to its popularity. The importance of designing image processing algorithms optimized for perceptual quality measures, as opposed to the MSE, has been long recognized [4]. Image coding algorithms that are optimized for perceptual distortion measures have been proposed by several authors and have become a part of image coding standards [9, 14, 10]. These algorithms use distortion measures such as those based on models of the human visual system (HVS), variants of the MSE such as the weighted MSE, or other empirical measures of quality.

Recent advances in image quality assessment (IQA) research has led to the emergence of powerful algorithms that include, among others, the SSIM Index [13], or Wang-Bovik Index,

the visual information fidelity criterion (VIF) [11], and the visual signal to noise ratio (VSNR) [2]. These IQA algorithms correlate better with the mean opinion scores (MOS) of subjective quality evaluation compared to measures currently used in image coding algorithms such as the just-noticeable distortion (JND) metric [8] and MSE [12]. The emergence of these IQA algorithms present the possibility of improved design of image coding algorithms.

In this paper, we analyze the relation between the quantization rate and the SSIM Index in the discrete cosine transform (DCT) domain. The SSIM Index is considered as opposed to other recent IQA algorithms since its performance is very competitive while being highly amenable to analysis. The DCT domain is considered since it is popularly used in transform coding of images and videos. A closed form relation between rate and SSIM Index for fixed rate uniform quantization is extremely difficult, if not impossible, to derive. Instead, we are able to derive upper and lower bounds on the SSIM Index as a function of quantization rate (under a high rate assumption). It is shown that the SSIM Index between the reference and quantized versions of the input falls within these bounds not only for Gaussian and Laplacian sources, but also for natural images. The usefulness of the bounds in a practical image coding scenario is demonstrated using a simple rate allocation example.

The SSIM Index is briefly reviewed followed by an overview of the discrete cosine transform (DCT) and uniform quantization. The rate versus SSIM Index problem is then formulated in the DCT domain. Bounds on the SSIM Index as a function of quantization rate are derived for Gaussian and Laplacian sources. The usefulness of the bounds in a practical image coding scenario is demonstrated using a bit allocation example.

2 Background

2.1 SSIM Index

The most general form of the structural similarity measure between two signal vectors $\mathbf{x}$ and $\mathbf{y}$ (both in $\mathbb{R}^N$) is

$$SSIM(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha [c(\mathbf{x}, \mathbf{y})]^\beta [s(\mathbf{x}, \mathbf{y})]^\gamma. \tag{1}$$

The term $l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$ compares the luminance of the signals, $c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$ compares the contrast of the signals, and $s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$ measures the structural correlation of the signals. The quantities μ_x, μ_y are the sample means of $\mathbf{x}$ and $\mathbf{y}$ respectively, σ_x^2, σ_y^2 are the sample variances of $\mathbf{x}$ and $\mathbf{y}$ respectively, and σ_{xy} is the sample cross-covariance between $\mathbf{x}$ and $\mathbf{y}$. The constants C_1, C_2, C_3 are used to stabilize the distortion measure for the case where the means and variances become small. The parameters $\alpha > 0$, $\beta > 0$, and $\gamma > 0$, are used to adjust the relative importance of the three components. We use the following simplified form of the SSIM Index in our work (where $\alpha = \beta = \gamma = 1$, and $C_3 = C_2/2$):

$$SSIM(\mathbf{x}, \mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left(\frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right). \tag{2}$$

In image quality assessment, pixel values from local blocks of the reference and distorted images constitute the vectors $\mathbf{x}$ and $\mathbf{y}$ respectively. The average of the SSIM values across

the image (also called mean SSIM or MSSIM) gives the final quality measure. The design philosophy of the SSIM Index is to acknowledge the fact that natural images are highly structured, and that the measure of structural correlation (between the reference and the distorted image) is important for deciding overall visual quality. Further, the SSIM Index measures quality locally and is able to capture local dissimilarities better, unlike global quality measures such as MSE (and hence PSNR). Though (2) has a form that is more complicated than MSE, it remains analytically tractable.

2.2   The Discrete Cosine Transform

The DCT is widely used in the transform coding of images and videos and is central to several popular image (JPEG) and video coding standards (MPEG-x) [10, 3]. Highly efficient software and hardware implementations of the DCT form the core of several of these standards. The DCT is popular due to its energy compaction property, combined with efficient implementations. These reasons motivate us to perform our analysis in the DCT domain. The DCT of a vector $\mathbf{x} \in \mathbb{R}^N$ is:

$$X(k) = \sum_{i=0}^{N-1} \lambda(k) cos \left(\frac{(2i+1)\pi k}{2N} \right) x(i), \tag{3}$$

where $\lambda(k) = \sqrt{\frac{1}{N}}$, if $k = 0$, $\lambda(k) = \sqrt{\frac{2}{N}}$ otherwise.

2.3   Uniform Quantization

Uniform quantization [1] is the earliest, simplest, and most common form of quantization. It is used in a range of audio, image, and video coding applications mainly due to its simplicity. While other forms of quantization are well studied [5], asymptotic analysis of the relation between rate and distortion (mean squared error) for fixed-rate uniform quantization of symmetric sources with infinite support was reported only as recently as 2001 by Hui and Neuhoff [6]. We use results from this work in our implementation.

The following notation is used in our analysis. The range of the granular region is denoted by $2L$, the number of quantization levels $N = 2^R$, where R is the quantization rate. The quantizer step size is denoted by $\Delta = \frac{2L}{N}$. The quantization levels are denoted by y_i ($0 \leq i \leq N-1$), with $y_i = -L + (i + 1/2)\Delta$. An interval in the granular region is denoted by $S_i = (y_i - \Delta/2, y_i + \Delta/2]$. In the sequel, we use the term rate and quantization step size interchangeably (for notational convenience) since they are related by $\Delta = \frac{2L}{2^R}$, where $2L$ and R are as defined above.

3   Problem Formulation

3.1   Measuring SSIM Index from DCT Coefficients

The SSIM Index in (2) is defined in the space domain. In the following, we derive a simple yet useful formula for measuring the SSIM Index between two vectors from their DCT coefficients. The DCT is a unitary transform and obeys Parseval's theorem. Using this property

and (3), the following relations between space domain mean, variance, cross correlation and the DCT coefficients are established:

$$\mu_x = \frac{\sum_{i=0}^{N-1} x(i)}{N} = \frac{X(0)}{\sqrt{N}}, \sigma_x^2 = \frac{\sum_{k=1}^{N-1} X(k)^2}{N-1}, \sigma_{xy} = \frac{\sum_{k=1}^{N-1} X(k)Y(k)}{N-1}. \tag{4}$$

Substituting the space domain mean, variance, and cross correlation terms in the definition of SSIM (2) with the expressions in (4),

$$SSIM(\mathbf{x}, \mathbf{y}) = \left(\frac{2\frac{X(0)Y(0)}{N} + C_1}{\frac{X(0)^2 + Y(0)^2}{N} + C_1} \right) \left(\frac{2\frac{\sum_{k=1}^{N-1} X(k)Y(k)}{N-1} + C_2}{\frac{\sum_{k=1}^{N-1} X(k)^2 + Y(k)^2}{N-1} + C_2} \right). \tag{5}$$

This expression can be particularly useful when performing quality assessment of JPEG compressed images without having to decompress the images to the space domain (for computing the SSIM Index from non-overlapping blocks). We use the DCT domain expression for the SSIM Index in the following analysis.

3.2   Relation between SSIM Index and Quantization Rate

The relation between SSIM and rate is derived under the high rate assumption and includes contributions only from the granular region. We assume that the DCT coefficients are independent [7], and that they are quantized at different step sizes Δ_i [10].

Let $\mathbf{X} = [X_0 \ X_1 \ \ldots \ X_{N-1}]^T$ denote a random vector composed of N DCT coefficients. In the sequel we assume that the elements $X_0, X_1, \ldots, X_{N-1}$ are independent and have a joint density $f_{\mathbf{X}}(\mathbf{x}) = f(x_0)f(x_1) \ldots f(x_{N-1})$. Each element of the random vector X_i is uniformly quantized at rate R_i. Under these assumptions, an interval in the joint granular region of the quantizers can be indexed by a vector $\mathbf{k} = [k_0, k_1, \ldots, k_{N-1}]$, where k_i varies between 0 and $2^{R_i} - 1$. The upper and lower limits of a quantization interval are denoted by U_{k_i} and L_{k_i} respectively. The vector $\mathbf{X}(\in S_{\mathbf{k}})$ is quantized by a point $Q(\mathbf{X}) = [Y_{k_0} \ Y_{k_1} \ \ldots \ Y_{k_{N-1}}]^T$. We ignore the contribution of the overload region to the average SSIM between $\mathbf{X}$ and $Q(\mathbf{X})$, and consider only the granular region. The average SSIM Index between $\mathbf{X}$ and $Q(\mathbf{X})$ is computed as

$$\mathbb{E}[SSIM(\mathbf{X}, Q(\mathbf{X}))] \approx \mathbb{E}[SSIM_{gran}(\mathbf{X}, Q(\mathbf{X}))]$$

$$= \sum_{\mathbf{k}} \int_{L_{k_0}}^{U_{k_0}} \frac{2x_0 Y_{k_0} + NC_1}{x_0^2 + Y_{k_0}^2 + NC_1} f(x_0)dx_0 \int_{L_{k_1}}^{U_{k_1}} \ldots \int_{L_{k_{N-1}}}^{U_{k_{N-1}}} \left(\frac{2\sum_{i=1}^{N-1} x_i Y_{k_i} + (N-1)C_2}{\sum_{i=1}^{N-1} x_i^2 + Y_{k_i}^2 + (N-1)C_2} \right) \tag{6}$$

$$f(x_1)f(x_2) \ldots f(x_{N-1})dx_1 dx_2 \ldots dx_{N-1}.$$

In practice, the most common DCT block size used in image and video coding applications is 8×8. The expression in (6) however, is quite formidable to evaluate and implement even for DCT block sizes as small as 2×2. Therefore, directly using (6) in a practical scenario appears extremely difficult, if not impossible. To make this problem tractable, we develop upper and lower bounds on (6). These bounds are shown to be accurate in estimating the

4

range of the average SSIM Index between the reference and quantized versions of a variety of sources. Further, it is also shown that the bounds are easier to implement and evaluate than an explicit solution to (6).

4   Bounds on the SSIM Index

In this section we present upper and lower bounds on the average SSIM Index as a function of quantization rate, evaluate these bounds for Gaussian and Laplacian sources, and discuss several properties of these bounds. We assume that the DCT coefficients X_i are independent, and each coefficient is quantized separately at step size Δ_i. The high resolution assumption is made, and only the contribution of the granular region is considered.

Theorem 4.1 *For a random vector* $\mathbf{X}$ *with independent components, the average SSIM Index (as defined in (5), (6)) between* $\mathbf{X}$ *and its uniformly quantized version* $Q(\mathbf{X}) = [Q_0(X_0), Q_1(X_1), \ldots, Q_{N-1}(X_{N-1})]^T$ *is bounded with probability p by*

$$\bar{M}\left(1 - \frac{\frac{\sum_{i=1}^{N-1}\Delta_i^2/12}{N-1}}{U + C_2}\right) \leq \mathbb{E}[SSIM(\mathbf{X}, Q(\mathbf{X}))] \leq \bar{M}\left(1 - \frac{\frac{\sum_{i=1}^{N-1}\Delta_i^2/12}{N-1}}{V + C_2}\right), \qquad (7)$$

where Δ_i is the step size assigned to quantizer Q_i to quantize random variable X_i, $\bar{M} = \mathbb{E}\left[\frac{2X_0Q(X_0)+NC_1}{X_0^2+Q(X_0)^2+NC_1}\right]$ is the average value of the contribution from the mean term, U, V are quantities defined below, $2L$ is the range of the granular region of the quantizer with the largest span, and C_2 is a stabilizing constant (from (5)).

The terms V and U for a given probability p are

$$\begin{aligned} V &= \sqrt{2}\sigma_v\text{erf}^{-1}(2p-1) + \mu_v, \\ U &= \sqrt{2}\sigma_u\text{erf}^{-1}(1-2p) + \mu_u \end{aligned} \qquad (8)$$

where $\sigma_v, \sigma_u, \mu_v, \mu_u$ are dependent on the source distribution.

Proof: Let vector $Q(\mathbf{X}) = [Y_{k_0}, Y_{k_1}, \ldots, Y_{k_{N-1}}]$ be a set of quantization levels corresponding to $\mathbf{X} = [X_0, X_1, \ldots, X_{N-1}]$.

$$SSIM(\mathbf{X}, Q(\mathbf{X})) = MS, \text{with}$$

$$M = \left(\frac{2\frac{X_0Y_{k_0}}{N} + C_1}{\frac{X(0)^2+Y_{k_0}^2}{N} + C_1}\right), S = \left(\frac{2\frac{\sum_{i=1}^{N-1}X_iY_{k_i}}{N-1} + C_2}{\frac{\sum_{i=1}^{N-1}X_i^2+Y_{k_i}^2}{N-1} + C_2}\right) \qquad (9)$$

where M corresponds to the mean term, and S corresponds to the structure term. It is easy to show that $M > 0$ (for natural images), and $-1 \leq S \leq 1$. Also, $1 - S \geq 0$. Now,

$$1 - S = 1 - \frac{2\frac{\sum_{k=1}^{N-1}X_iY_{k_i}}{N-1} + C_2}{\frac{\sum_{k=1}^{N-1}X_i^2+Y_{k_i}^2}{N-1} + C_2} = \frac{\frac{\sum_{i=1}^{N-1}(X_i-Y_{k_i})^2}{N-1}}{\frac{\sum_{i=1}^{N-1}X_i^2+Y_{k_i}^2}{N-1} + C_2}. \qquad (10)$$

5

To simplify the denominator in the above equation, a variable $Z = \frac{\sum_{i=1}^{N-1} X_i^2}{N-1}$ is introduced. Since N is large for the most interesting case of an 8×8 DCT, the distribution of Z can be approximated well by a Gaussian distribution due to the central limit theorem. For a specified probability p, choose V and U such that $P(Z \leq V - L^2) = p$, $P(Z > U) = p$. The term L^2 is subtracted from V since it is larger than the highest value that any of quantizer levels Y_{k_i} can take. Since Z is Gaussian, it follows that V, U that satisfy the probability p is given by (8), where the first and second moments of Z are computed based on the source distribution. Since $M > 0$, the following bound holds with probability p.

$$M \frac{\frac{\sum_{i=1}^{N-1}(X_i - Y_{k_i})^2}{N-1}}{V + C_2} \leq M(1 - S) \leq M \frac{\frac{\sum_{i=1}^{N-1}(X_i - Y_{k_i})^2}{N-1}}{U + C_2}. \tag{11}$$

Applying the expectation operator, using the high-rate uniform quantization result $MSE = \frac{\Delta^2}{12}$ [1, 5], and the independence assumption

$$\mathbb{E}[M][\frac{\frac{\sum_{i=1}^{N-1} \Delta_i^2/12}{N-1}}{V + C_2}] \leq \mathbb{E}[M] - \mathbb{E}[MS] \leq \mathbb{E}[M][\frac{\frac{\sum_{i=1}^{N-1} \Delta_i^2/12}{N-1}}{U + C_2}]. \tag{12}$$

Replacing $\mathbb{E}[M]$ with $\bar{M}$ and rearranging terms,

$$\bar{M}\left(1 - \frac{\frac{\sum_{i=1}^{N-1} \Delta_i^2/12}{N-1}}{U + C_2}\right) \leq \mathbb{E}[SSIM(\mathbf{X}, Q(\mathbf{X}))] \leq \bar{M}\left(1 - \frac{\frac{\sum_{i=1}^{N-1} \Delta_i^2/12}{N-1}}{V + C_2}\right) \square \tag{13}$$

The terms $\bar{M}, \sigma_v, \sigma_u, \mu_v, \mu_u$ are presented next for Gaussian and Laplacian sources. These two sources are considered as they are commonly used to model DCT coefficients [7]. The expressions for these bounds can be very easily implemented for these sources for any DCT block size. Most importantly, we show that the bounds are indeed accurate not only for these sources, but also for a large set of natural images.

In the following, the distribution of X_0 is assumed to have zero mean mainly to simplify notation. The essence of these results is the same irrespective of the mean. Their proofs are omitted for brevity.

4.1 Gaussian Source

If the DC coefficient X_0 is Gaussian distributed with zero mean and variance σ_0^2, the expression for $\bar{M}$ is given by

$$\bar{M} \approx \frac{1}{\sqrt{2\pi}\sigma_0} \sum_{k_0=0}^{2^{R_0}-1} Y_{k_0} e^{\frac{Y_{k_0}^2 + 2\sigma_0^2 NC_1}{2\sigma_0^2}} \left[E_1\left(\frac{L_{k_0}^2 + Y_{k_0}^2 + 2\sigma_0^2 NC_1}{2\sigma_0^2}\right) - E_1\left(\frac{U_{k_0}^2 + Y_{k_0}^2 + 2\sigma_0^2 NC_1}{2\sigma_0^2}\right)\right] \tag{14}$$

where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$ $(x > 0)$, is the exponential integral. The expression for $\bar{M}$ is an approximation in this case since we consider the contribution of only one term in the numerator.

If the AC coefficients $X_1, X_2, \ldots, X_{N-1}$ are independent and Gaussian distributed with zero mean and variance $\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N-1}^2$ respectively,

$$\sigma_v^2 = \frac{\sum_{i=1}^{N-1} 2\sigma_i^4}{(N-1)^2}, \mu_v = \frac{\sum_{i=1}^{N-1} \sigma_i^2}{(N-1)} + L^2, \sigma_u^2 = \frac{\sum_{i=1}^{N-1} 2\sigma_i^4}{(N-1)^2}, \mu_u = \frac{\sum_{i=1}^{N-1} \sigma_i^2}{(N-1)} \tag{15}$$

are a conservative set of parameters that satisfy the bounds.

4.2 Laplacian Source

If the DC coefficient X_0 is Laplacian distributed with zero mean and variance σ_0^2, the expression for $\bar{M}$ is a combination of three terms depending on the values of the upper and lower limits of the interval S_{k_0}. Suppose that there are N_1 intervals corresponding to Case 1 ($L_{k_0} > 0, U_{k_0} > 0$), N_2 intervals in Case 2 ($L_{k_0} \leq 0, U_{k_0} \leq 0$), and N_3 intervals in Case 3 ($L_{k_0} < 0, U_{k_0} > 0$), with $N_1 + N_2 + N_3 = 2^{R_0}$. Each case is evaluated as follows.

Case 1: $L_{k_0} > 0, U_{k_0} > 0$

$$T_1 = \frac{1}{\sqrt{2}\sigma_0} \sum_{k_0=0}^{N_1-1} Y_{k_0} \left\{ e^{iab}[E_1(a(L_{k_0} + ib)) - E_1(a(U_{k_0} + ib))] + e^{-iab}[E_1(a(L_{k_0} - ib)) - \right.$$

$$\left. E_1(a(U_{k_0} - ib))] \right\} + i\frac{NC_1}{\sqrt{b}} \left\{ e^{iab}[E_1(a(L_{k_0} + ib)) - E_1(a(U_{k_0} + ib))] - e^{-iab}[E_1(a(L_{k_0} - ib)) - \right.$$

$$\left. E_1(a(U_{k_0} - ib))] \right\}, \tag{16}$$

where $a = \frac{\sqrt{2}}{\sigma_0}, b = \sqrt{Y_{k_0}^2 + NC_1}, i = \sqrt{-1}, E_1(x)$ is the exponential integral.

Case 2: $L_{k_0} \leq 0, U_{k_0} \leq 0$

$$T_2 = \frac{1}{\sqrt{2}\sigma_0} \sum_{k_0=0}^{N_2-1} Y_{k_0} \left\{ e^{-iab}[E_1(-a(L_{k_0} + ib)) - E_1(-a(U_{k_0} + ib))] + e^{iab}(E_1[-a(L_{k_0} - ib)) - \right.$$

$$\left. E_1(-a(U_{k_0} - ib))] \right\} + i\frac{NC_1}{\sqrt{b}} \left\{ e^{-iab}[E_1(-a(L_{k_0} + ib)) - E_1(-a(U_{k_0} + ib))] - \right.$$

$$\left. e^{iab}[E_1(-a(L_{k_0} - ib)) - E_1(-a(U_{k_0} - ib))] \right\}, \tag{17}$$

where $a = \frac{\sqrt{2}}{\sigma_0}, b = \sqrt{Y_{k_0}^2 + NC_1}, i = \sqrt{-1}, E_1(x)$ is the exponential integral.

Case 3: $L_{k_0} < 0, U_{k_0} > 0$

$$T_3 = T_1 + T_2 \tag{18}$$

with T_1 as above evaluated over the interval $(0, U_{k_0}]$ and T_2 also as above, evaluated over $(L_{k_0}, 0]$.

$$\bar{M} = T_1 + T_2 + T_3. \tag{19}$$

If the AC coefficients $X_1, X_2, \ldots, X_{N-1}$ are independent and Laplacian distributed with zero mean and variance $\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N-1}^2$ respectively,

$$\sigma_v^2 = \frac{\sum_{i=1}^{N-1} 5\sigma_i^4}{(N-1)^2}, \mu_v = \frac{\sum_{i=1}^{N-1} \sigma_i^2}{(N-1)} + L^2, \sigma_u^2 = \frac{\sum_{i=1}^{N-1} 5\sigma_i^4}{(N-1)^2}, \mu_u = \frac{\sum_{i=1}^{N-1} \sigma_i^2}{(N-1)} \tag{20}$$

are a conservative set of parameters that satisfy the bounds.

4.3   Properties of the Bounds

The bounds in (7) possess several useful properties. (a) The terms $(\bar{M}, \sigma_v, \sigma_u, \mu_v, \mu_u)$ can be easily evaluated for several commonly used unbounded source types (as shown in the previous subsections). This property makes the bounds tractable when compared to (6). (b) In practice, different DCT coefficients are quantized at different rates in order to optimally allocate bits. The bounds hold for any combination of rates, thereby making them attractive in practical rate allocation problems. (c) From the expression for the bounds, we see that they can be implemented efficiently and easily (even for the complex looking Laplacian case). Note that the second term in the bound involves only N summation and division operations. This property could be very useful if these bounds were to be used in real-time codec implementations. This property also allows for fast computation at any practical DCT block size. (d) These bounds can easily be extended to the SSIM Index's predecessor - the universal image quality Index (UQI).

A point to note is that though the analysis considers a 1-D DCT, it is easy to show that the results carry over to the 2-D DCT case. The 2-D DCT obeys the Parseval's theorem, and the relation between the space domain and DCT domain means and inner products also hold.

5   Results

In this section, two representative results are presented. Rate versus SSIM Index plots for Gaussian and Laplacian sources are presented first, followed by a bit-allocation example of image sources.

5.1   Gaussian and Laplacian Sources

As a representative visual example, the case of equal rate allocation is considered for a set of 64 independent and identically distributed sources. The plot of the bounds, and the true SSIM Indices are shown in Fig. 1. For these plots, the probability p of the bounds being satisfied is set to 0.9. Since the values V, U are calculated using conservative estimates, it was found that the bounds hold with a much higher probability (almost 1) in practice. It can be observed from the plots that the bounds are accurate in the high rate regime.

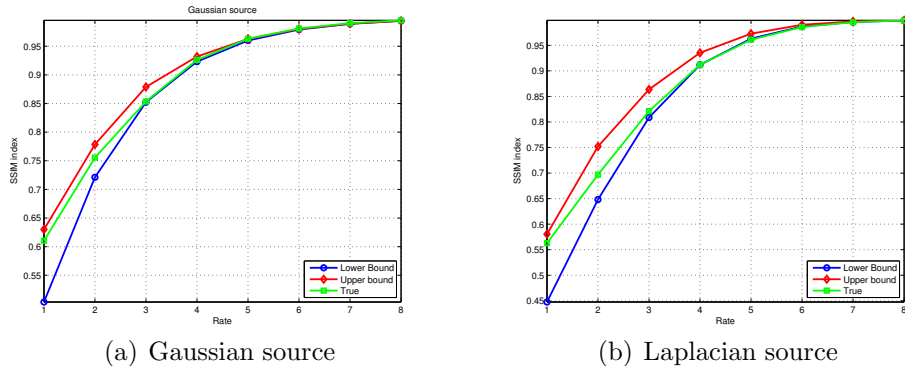(a) Gaussian source   (b) Laplacian source

Figure 1: Plots of rate versus SSIM Index for uniform quantization of 64 random variables that have zero mean and variance 100. 1(a) Gaussian source. 1(b) Laplacian source.

5.2 Bit-allocation Example

Consider the following rate allocation problem, and the associated constraints. Suppose that a bit budget of 128 bits is to be allocated to the 64 DCT coefficients. To make this problem tractable, the following constraints are introduced. The DCT coefficients are divided into four groups, each containing 16 coefficients. Further, the first group is assumed to contain the most important coefficients, the next group to contain the next most important coefficients, and so on. Finally, the more important group is always assigned bits greater than or equal to the number of bits assigned to the group immediately lower in importance. Though this setup is simple, it is a fair reflection of a true coding scenario. Under these assumptions, four combinations are possible - (5, 1, 1, 1), (4, 2, 1, 1), (3, 3, 1, 1), and (2, 2, 2, 2). The problem is to find the rate combination that results in the highest SSIM Index of the quantized image. The results of allocating these bit combinations to an image are shown in Fig. 5.2. Also shown are the predicted SSIM Indices using the average of proposed upper and lower bounds for Laplacian sources (with $p = 0.9$), and true SSIM Indices. It can be seen that the bounds help make an accurate prediction of the SSIM Index.

6 Conclusions

In this paper, we presented bounds on the SSIM Index as a function of quantization rate for fixed-rate uniform quantization. The proposed bounds make use of a well-known relation between MSE and quantization rate for fixed rate uniform quantization under the high rate assumption. We have demonstrated the strength of the proposed bounds for Gaussian and Laplacian sources and their usefulness in a practical image coding scenario. Through these results, we have taken a step in the path of designing perceptually optimal image coding algorithms, and more generally in designing perceptually optimal image processing algorithms that leverage the strength of emerging IQA algorithms.

9

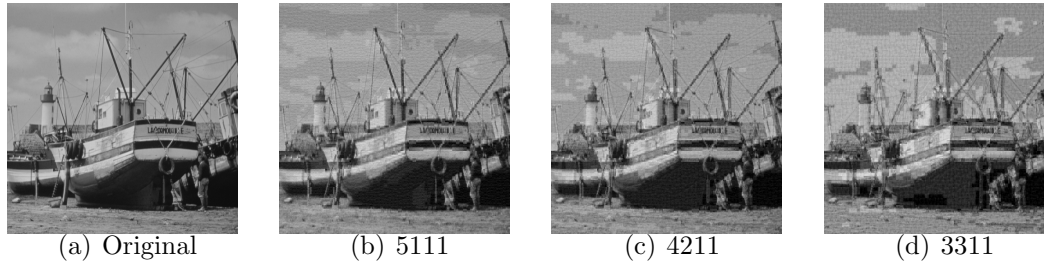| (a) Original | (b) 5111 | (c) 4211 | (d) 3311 |

Figure 2: *Rate allocation example*: 2(a) The original Boats image. 2(b) Original quantized using the 5, 1, 1, 1 rate profile. SSIM Index = 0.7743, Laplacian estimate of SSIM Index = 0.7755. 2(c) Original quantized using the 4, 2, 1, 1, rate profile. SSIM Index = 0.7551, Laplacian estimate of SSIM Index = 0.7584. 2(d) Original quantized using the 3, 3, 1, 1 rate profile. SSIM Index = 0.6689, Laplacian estimate of SSIM Index = 0.6583.

References

[1] W. R. Bennett, "Spectra of Quantized signals," *Bell Systems Technical Journal*, vol. 27, pp. 446–472, July 1948.

[2] D. M. Chandler and S. S. Hemami, "VSNR: A Wavelet-based Visual Signal-to-Noise Ratio for Natural Images," *IEEE Trans. Image Processing*, vol. 16, pp. 2284–2298, 2007.

[3] C. Fogg, D. J. LeGall, J. L. Mitchell, and W. B. Pennebaker, *MPEG Video Compression Standard*. Springer, 1996.

[4] B. Girod, "What's Wrong with Mean-squared Error?" in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993, pp. 207–220.

[5] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[6] D. Hui and D. L. Neuhoff, "Asymptotic Analysis of Optimal Fixed-Rate Uniform Scalar Quantization," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 957–977, Mar. 2001.

[7] E. Y. Lam and J. W. Goodman, "A Mathematical Analysis of the DCT Coefficient Distributions for Images," *IEEE Trans. Image Processing*, vol. 9, no. 10, pp. 1661–1666, Oct. 2000.

[8] J. Lubin, *Visual Models for Target Detection and Recognition*. World Scientific, 1995, ch. 10, pp. 245–283.

[9] J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," *IEEE Trans. Inform. Theory*, vol. 20, pp. 525–536, Jul. 1974.

[10] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*. New York: Van Nostrand Reinhold, 1993.

[11] H. R. Sheikh and A. C. Bovik, "Image Information and Visual Quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[12] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Trans. Image Processing*, vol. 15, no. 11, pp. 3440 – 3451, Nov 2006.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[14] A. Watson, "Visually Optimal DCT Quantization Matrices for Individual Images," in *Data Compression Conference, 1993. DCC '93.*, 30 March-2 April 1993, pp. 178–187.