

DCT Quantization Noise in Compressed Images

Mark A. Robertson and Robert L. Stevenson

Manuscript received ———.

Robert L. Stevenson is with the Laboratory for Image and Signal Analysis (LISA) in the Electrical Engineering Department at the University of Notre Dame. Mark A. Robertson was with the LISA, and is currently with the Air Force Research Laboratory. Portions of this work appear in [1, 2].

This work was supported in part by Indiana's 21st Century Fund and by Sun Microsystems.

Abstract

In lossy image compression schemes utilizing the discrete cosine transform (DCT), quantization of the DCT coefficients introduces error in the image representation and a loss of signal information. At high compression ratios, this introduced error produces visually undesirable compression artifacts that can dramatically lower the perceived quality of a particular image. This paper provides a spatial domain model of the quantization error based on a statistical noise model of the error introduced when quantizing the DCT coefficients. The resulting theoretically derived spatial domain quantization noise model shows that in general the compression noise in the spatial domain is both correlated and spatially varying. This provides some justification to many of the ad hoc artifact removal filters that have been proposed. More importantly, the proposed noise model can be incorporated in a post-processing algorithm that correctly incorporates the spatial correction of the quantizer error. Experimental results demonstrate the effectiveness of this approach.

Keywords

DCT Quantization Noise, Image Restoration, Compressed Image Post-Processing, DCT Coefficient Distributions

I. INTRODUCTION

Artifacts in compressed images are well-known phenomena, both in the scientific and consumer communities. In order to compensate for these artifacts, one must have a thorough understanding of them. This paper derives statistical descriptions of the quantization noise that, in addition to providing important insight into the origins of compression noise, can also be readily applied in a post processing algorithm to remove the artifacts.

Blocking is one of the most common forms of compression artifact, and many attribute its existence to the inability of the block DCT to include correlations between blocks. While there is truth to this, the blocking phenomenon can also be explained by considering the compression of a single block only, which is the approach taken here: Only a single block of the image to which the DCT is applied is examined, and results are derived for that block. As will be seen, results for a single block prove sufficient to explain blocking artifacts, as well as compression noise in general.

Section II provides the basics of the DCT and establishes notation used in the remainder of the paper. A review of other explicitly or implicitly proposed noise models, including both their advantages and limitations, is presented in Section III to provide motivation for this work. Sections IV and V derive a statistical description of the noise introduced by quantization of

DCT coefficients. In particular, Equation (12) gives the Gaussian distribution of the spatial-domain quantization error, with autocovariance given by (9). The spatial-domain autocovariance depends on frequency-domain characteristics of the image, as described by Equations (13) and (19). A discussion of the work is provided in Section VI, where an explanation for the existence of blocking artifacts is proposed, including where one can expect their presence or absence. Section VI also includes experimental results that demonstrate the principles put forth in this paper. An example application of the work is presented in Section VII, where the DCT quantization noise model is used in a post-processing algorithm to remove compression artifacts. Concluding remarks are given in Section VIII.

II. THE DISCRETE COSINE TRANSFORM

In one dimension, the signal to be transformed is denoted $z[n]$, $0 \leq n < N$, which is also represented by the length- N column vector \mathbf{z} . The elements of the DCT matrix $\mathbf{H} = \{\mathbf{H}[k, n]\}$ are defined as

$$\mathbf{H}[k, n] = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0, 0 \leq n < N, \\ \sqrt{\frac{2}{N}} \cos \left[\frac{\pi(2n+1)k}{2N} \right], & 1 \leq k < N, \\ 0 & 0 \leq n < N. \end{cases} \quad (1)$$

The DCT of \mathbf{z} is then $\mathbf{y} = \mathbf{H}\mathbf{z}$. Since the DCT is a real unitary transform, $\mathbf{H}^{-1} = \mathbf{H}^t$ and the inverse DCT (IDCT) is described by $\mathbf{z} = \mathbf{H}^t\mathbf{y}$. Letting \mathbf{h}_k^t denote the k^{th} row of \mathbf{H} ,

$$\mathbf{z} = \sum_{k=0}^{N-1} \mathbf{h}_k \mathbf{y}[k]. \quad (2)$$

Equation (2) gives a series representation of the vector \mathbf{z} in terms of the DCT basis vectors \mathbf{h}_k and the DCT coefficients $\mathbf{y}[k]$. Any one-dimensional (1-D) signal of length N can be represented by a sum of these N basis vectors of different frequencies, where each basis vector in the sum is scaled by the DCT coefficient $\mathbf{y}[k]$.

In a DCT-based compression scheme, the DCT coefficients are quantized rather than the actual signal \mathbf{z} . The quantized DCT coefficients are $\mathbf{y}_q = Q[\mathbf{y}]$, where $Q[\cdot]$ is the quantization operator. Quantization is a non-linear operation that results in a loss of information; only scalar quantization is considered here, where each element of \mathbf{y} is individually quantized. Scalar quantization is a many-to-one mapping that transforms intervals of real numbers $[q_i^k, q_{i+1}^k)$ to single real numbers. The superscript “ k ” accounts for the possibility of different quantization intervals

for different frequency coefficients, and the subscript “ i ” indicates the i^{th} quantization level. Transform coefficients that are in these intervals are typically mapped to the midpoint of the interval, so that $y_q[k] = \frac{1}{2}(q_i^k + q_{i+1}^k)$ for $q_i^k \leq y[k] < q_{i+1}^k$.

The recovered signal \mathbf{z}_q is found by performing the inverse DCT on the quantized frequency values, $\mathbf{z}_q = \mathbf{H}^t \mathbf{y}_q$. Two quantities of interest in this paper are the quantization errors in both the spatial and the frequency domains. The vector $\mathbf{e}_z = \mathbf{z}_q - \mathbf{z}$ will represent the error in the spatial domain, and $\mathbf{e}_y = \mathbf{y}_q - \mathbf{y}$ will represent the error in the frequency domain. Note that the quantization error in the spatial domain can be expressed as

$$\mathbf{e}_z = \mathbf{H}^t [\mathbf{y}_q - \mathbf{y}] = \sum_{k=0}^{N-1} \mathbf{h}_k (y_q[k] - y[k]). \quad (3)$$

Equation (3) gives the basis representation of the quantization error: The quantization error is the sum of the errors of each frequency coefficient multiplied by the corresponding DCT basis vector.

The two-dimensional (2-D) DCT is simply a separable extension of the DCT in one dimension. Here, the matrix \mathbf{Z} represents the 2-D image block, with \mathbf{Y} , \mathbf{Z}_q , \mathbf{Y}_q , \mathbf{E}_Z and \mathbf{E}_Y defined analogously to the 1-D case.

As for the 1-D case, one can write a basis representation for $N \times N$ image blocks. If the $N \times N$ basis images are defined as $\mathbf{H}_{k,l} = \mathbf{h}_k \mathbf{h}_l^t$, then

$$\mathbf{Z} = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{H}_{k,l} Y[k, l]. \quad (4)$$

Figures showing the basis vectors and images defined by (2) and (4) can be found in many standard image processing textbooks.

Quantization of the 2-D DCT coefficients is performed as for the 1-D case, with a resulting spatial-domain quantization error of

$$\mathbf{E}_Z = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{H}_{k,l} (Y_q[k, l] - Y[k, l]), \quad (5)$$

which gives the basis-image representation of the quantization error: The quantization error is the sum of the errors of each transform coefficient multiplied by each of the 2-D DCT basis images. This way of looking at the quantization error in compressed images has also been considered in [3, 4].

On some occasions it will be convenient to use vector notation to represent images. When using this vector notation, the vector \mathbf{z} of length N^2 will represent the image block \mathbf{Z} , where $z[lN + k] = Z[k, l]$. This is equivalent to stacking the individual columns of \mathbf{Z} into a single vector. The matrix \mathcal{H} is then defined as the Kronecker product of \mathbf{H} with itself, $\mathcal{H} = \mathbf{H} \otimes \mathbf{H}$. The $N^2 \times N^2$ matrix \mathcal{H} is the 2-D DCT matrix for image blocks in vector format. Denoting the various image matrices from before with this stacked-column vector notation, the results can be summarized as $\mathbf{y} = \mathcal{H}\mathbf{z}$, $\mathbf{z} = \mathcal{H}^t\mathbf{y}$, and $\mathbf{z}_q - \mathbf{z} = \mathcal{H}^t(\mathbf{y}_q - \mathbf{y})$.

From (3) and (5), the spatial-domain error introduced by image compression is dependent on two key factors: The DCT basis vectors, and the frequency-domain error introduced by quantization of the DCT coefficients. Statistics of the spatial-domain quantization noise given statistics of the frequency-domain quantization noise will be examined in Section IV, followed by statistics of the noise in the frequency domain in Section V. First, however, other proposed models for image compression noise will be examined.

III. REVIEW OF OTHER NOISE MODELS

While the quantization error is a deterministic function of the input image, in many practical applications once the quantized signal \mathbf{y}_q is calculated, the input image \mathbf{y} is discarded, and thus explicit information about the quantization error is lost. A commonly used theoretic tool for modeling the error signal is to treat it as a random quantity [5]. Treating the error as random provides an understanding of how the error behaves, including how much the error will vary at different pixel locations and the spatial correlation between errors. Such understanding further provides the theoretical framework for formulating effective schemes for alleviating the error. This statistically-based model of the error signal is referred to as a quantization noise model since the error signal represents unwanted information in the resulting image representation.

Characterizing DCT quantization noise is an important task if one wishes to attempt to remove the noise, and many of the noise models in the literature have been proposed as part of post-processing schemes. In the Bayesian post-processing scheme of [6], the authors make use of a quantization constraint set (QCS), which is the set of all images that would re-compress to agree with the (quantized) received data. The QCS constraint is applied in the algorithm by clipping reconstructed DCT coefficients to force them to lie within the quantization limits defined by the compression. Use of the QCS is equivalent to assuming a uniform distribution of the DCT

coefficients—DCT coefficients of reconstructed images can lie anywhere within the quantization levels, with no preference for exactly where as long as they do not go outside the quantization limits.

Although not explicitly considering noise or probabilistic methods, the projection onto convex sets (POCS) post-processing methods [7] implicitly use the same noise model as in [6]. One of the convex sets used in the POCS algorithm in [7] is the very QCS used in [6]. Thus, although based on very different theories, the Bayesian and POCS methods in [6, 7] assume identical noise models.

One criticism of the QCS is that it is often too big, i.e., the variety of images that lie within the QCS is too large, and the QCS fails to model the actual compression noise adequately. Figure 1 demonstrates this concept, where the limitations of the QCS are apparent. While both Figs. 1(c) and 1(d) lie within the QCS, they are both poor visual estimates of the original image in Figure 1(a). The reason for this inadequacy can be seen by considering the following: Many images (or portions of images) are quite often smooth, i.e., they have most of their energy concentrated at low frequencies, with little energy at high frequencies. This situation provides an example where the uniform noise model fails: If a quantized high-frequency coefficient of a “smooth” image is observed to be zero, then modeling the original coefficient as varying uniformly throughout the entire quantization interval will not be accurate. Rather, the model should incorporate the idea that the original coefficient was quite close to zero, and in fact did not vary far from zero.

An interesting extension of the POCS post-processing theory is given in [8], where the authors address the limitation just described for the QCS. The authors introduce a narrow quantization constraint set (NQCS), which is a subset of the QCS defined to make the constraint set “tighter.” The NQCS essentially assumes the same uniform model as in [6,7], but the interval of each DCT coefficient is artificially made smaller.

Pointing out limitations of the QCS, Meier et al. [9] chose to adopt a compression noise model exclusively in the spatial domain. They assumed independent and identically-distributed (IID) Gaussian random variables for the quantization noise. Mateos et al. [10] also use IID Gaussian noise, but only for the block boundary pixels that they process. The noise models of these probabilistic methods are mathematically equivalent to a constrained least squares framework [11].

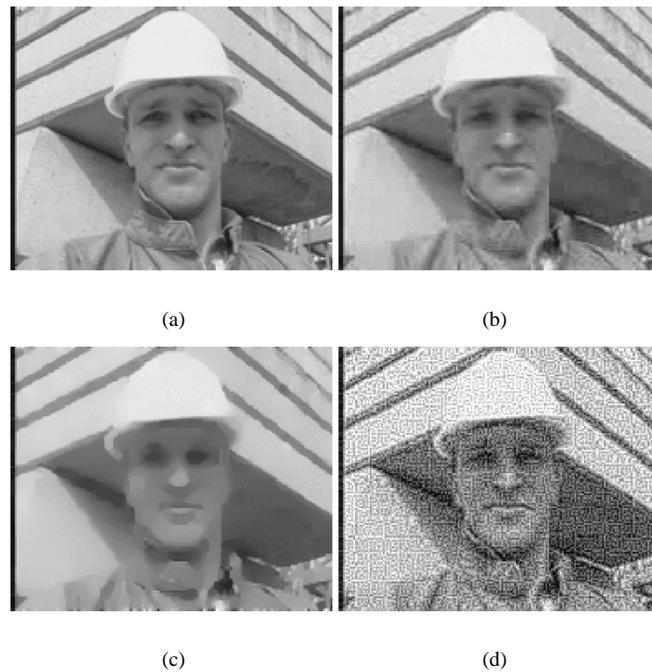


Fig. 1. Different images that each lie in the same quantization constraint set of a P-frame of the *foreman* sequence: (a) the original image; (b) compressed image with standard de-compression; (c) compressed image that has been smoothed subject to quantization constraints; and (d) compressed image that has been sharpened subject to quantization constraints.

These methods just mentioned bypass the problem of the QCS methods: The actual amount of assumed noise can be controlled by adjustment of the variance of the Gaussian noise, which can effectively prevent over-smoothing in a post-processing scheme.¹ However, since these methods do not consider the actual compression process, they do not accurately predict the quantization noise. Furthermore, the IID assumption cannot actually model the spatial correlation present in the quantization noise.

Post-processing schemes that use the QCS or NQCS do not have explicit spatial-domain noise models, but rather work in the frequency domain. The main advantage of using a quantization constraint set is that it explicitly considers the method of compression, and hence provides a noise model closer to reality than simple IID Gaussian noise in the spatial domain. However, as mentioned earlier, the QCS methods have a severe limitation in that the constraint sets are too large. An additional disadvantage of QCS based methods is that the resulting post-processing al-

¹Equivalently, a regularization parameter can be varied to control the amount of smoothing to perform, where the regularization parameter affects the *a priori* smoothness assumed for the image.

gorithms generally require multiple transformations between the frequency-domain and spatial-domain representations of the image during post-processing, which is a computationally expensive task,

Yang et al. [12] also use a noise model in the spatial domain, but do not derive actual quantization noise statistics, choosing rather to work with a training procedure to determine noise characteristics experimentally. They do *not* assume identical distributions—based on a training procedure, their algorithm determines noise vectors that approximate the shape of the compression noise, which often has large magnitude near block boundaries. This effectively suggests that quantization noise is higher near block boundaries than in block interiors.

DCT quantization noise has also been studied in contexts other than post-processing. Peterson et al. [4, 13] studied the effects of quantization noise, but they were primarily concerned with the visibility of this noise to human observers and do not consider the statistics of the noise itself. Yovanof and Liu [14] examine DCT spatial-domain quantization noise experimentally, but do not consider the possibility of spatially-varying noise or correlated noise.

Limitations of a spatially-invariant quantization noise model are shown in Fig. 2. The figure plots the mean squared pixel error for each 8×8 block of the 320×288 *claire* sequence. The sequence was compressed according to the H.263 standard [15] with a constant quantization parameter (QP) of 12, using only I-frames. If the quantization noise were in fact spatially invariant, the averaged noise for each 8×8 block would be approximately flat. This is clearly not the case for the plot in Fig. 2, which actually suggests that the quantization error is higher for pixels near block boundaries, and especially high at block corners.

Two important conclusions can be drawn from this section’s review of DCT quantization noise models. First, those frequency-domain noise models that make use of the QCS can be inadequate due to the potentially large sets that result. In addition, the spatial-domain noise models were limited as well—the IID assumption does not accurately model actual compression noise. In the next section a spatial domain quantization noise model is proposed that actually models the compression noise and which can be effectively used within a post-processing algorithm.

IV. QUANTIZATION NOISE STATISTICS

One reason that the DCT is utilized in image compression is its ability to approximately decorrelate the type of signals found in image data [16]. That is, while the typical image signal

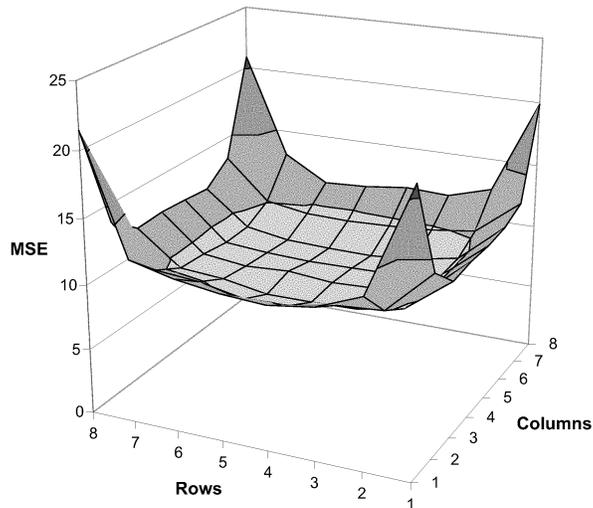


Fig. 2. Average quantization error variance for each block of the 320×272 *claire* sequence compressed using I-frames only, with H.263 quantization parameter 12.

is highly correlated, the DCT coefficients $y[k]$ are approximately uncorrelated. This property can be used to approximate the DCT coefficients $y[k]$ as uncorrelated random variables. For relatively high-rate situations (rather small quantization bin sizes), it is well established that the quantization errors are also uncorrelated [17, 18]. For lower-rate situations it is also possible to show that, given some mild conditions, the quantization errors in the DCT domain are uncorrelated.

For the reasons just outlined, the DCT-domain quantization errors are assumed to be uncorrelated random variables. Considering the one-dimensional case first, the covariance matrix $\mathbf{K}_{e_y} = E[(\mathbf{y}_q - \mathbf{y})(\mathbf{y}_q - \mathbf{y})^t | \mathbf{y}_q]$ is then diagonal, with its N non-zero elements equal to the quantization noise variances of the individual frequency-domain coefficients, $\sigma_{e_y}^2[k]$. Note that the expectation for \mathbf{K}_{e_y} is conditioned on \mathbf{y}_q , implying that the quantity describes the expected error after receiving the compressed image data. For notational convenience, the conditioning on the received data will not be explicitly written in the remainder of this paper (for \mathbf{y}_q or \mathbf{z}_q), but will rather be implicitly assumed.

Suppose for the moment that \mathbf{K}_{e_y} is known; it is then simple to determine statistics of the error in the spatial domain. Of primary interest here is the autocovariance

$$\mathbf{K}_{e_z} = E[(\mathbf{z}_q - \mathbf{z})(\mathbf{z}_q - \mathbf{z})^t] = \mathbf{H}^t \mathbf{K}_{e_y} \mathbf{H}. \quad (6)$$

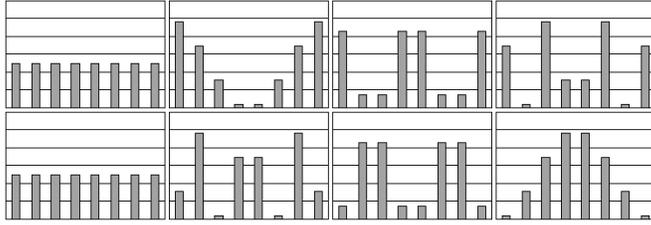


Fig. 3. One-dimensional basis vectors of the spatial-domain quantization error variance for sequences of length 8. The horizontal axis of each plot corresponds to spatial location n and the frequencies start with $k = 0$ for the top left and continue to $k = 7$ for the bottom right.

Equation (6) includes information about the autocovariance of the spatial-domain error vector, but another quantity of interest is the variance of the individual spatial-domain errors. The variance of $e_z[n]$ is found as $\sigma_{e_z}^2[n] = \mathbf{K}_{e_z}[n, n]$, or in summation notation

$$\sigma_{e_z}^2[n] = \sum_{k=0}^{N-1} H^2[k, n] \mathbf{K}_{e_y}[k, k] = \sum_{k=0}^{N-1} H^2[k, n] \sigma_{e_y}^2[k], \quad (7)$$

where $H^2[k, n]$ is the square of the $(k, n)^{th}$ element of \mathbf{H} . Based on this equation, a new transformation matrix \mathbf{M} is defined, with elements $M[k, n] = H^2[k, n]$. Letting \mathbf{m}_k^t denote the k^{th} row of \mathbf{M} ,

$$\sigma_{e_z}^2 = \sum_{k=0}^{N-1} \mathbf{m}_k \sigma_{e_y}^2[k]. \quad (8)$$

Here, $\sigma_{e_z}^2$ is the vector representing the error variance of \mathbf{z}_q , which is also equivalent to the diagonal of \mathbf{K}_{e_z} . The vectors \mathbf{m}_k form a basis for the error variance of \mathbf{z}_q : The error variance can be found by summing each error variance basis vector scaled by the corresponding frequency-domain error variance. Error variance basis vectors for sequences of length 8 are shown in Fig. 3. Transform-domain coefficients with large quantization error variances cause their corresponding basis vectors to contribute more to the overall spatial-domain error variances. In general, this leads to spatial-domain error variances that vary with n , the spatial location. Also, note from (6) that in general the spatial-domain quantization noise is correlated.

To represent the autocovariance of the noise in two dimensions, it is helpful to use the stacked-column notation introduced earlier, where the random vectors \mathbf{z} , \mathbf{y} , \mathbf{z}_q , \mathbf{y}_q , \mathbf{e}_z and \mathbf{e}_y are formed by stacking the columns of their respective random images. As before, assume for the moment that the covariance matrix of the 2-D frequency-domain quantization noise is a known diagonal matrix \mathbf{K}_{e_y} , whose N^2 diagonal elements are taken from the individual frequency-domain error

variances $\sigma_{\mathbf{E}_Y}^2[k, l]$. The covariance matrix of the spatial-domain error can then be written as

$$\mathbf{K}_{\mathbf{e}_z} = E \left[(\mathbf{z}_q - \mathbf{z})(\mathbf{z}_q - \mathbf{z})^t \right] = \mathcal{H}^t \mathbf{K}_{\mathbf{e}_y} \mathcal{H}. \quad (9)$$

The $N^2 \times N^2$ matrix $\mathbf{K}_{\mathbf{e}_z}$ contains autocovariance information of the spatial-domain error image, but the individual error variances at the N^2 locations are also of interest. These individual variances are the diagonal elements of $\mathbf{K}_{\mathbf{e}_z}$. Switching back to 2-D notation, they can be written in a form similar to (7),

$$\sigma_{\mathbf{E}_Z}^2[m, n] = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} H^2[l, m] H^2[k, n] \sigma_{\mathbf{E}_Y}^2[k, l]. \quad (10)$$

Choi and Kim also include this equation in their paper on artifact reduction [19], but make no mention of the error covariance as described here.²

The basis-image representation for the quantization error variance of the DCT in two dimensions can be found by defining a set of images $\mathbf{M}_{k,l} = \mathbf{m}_k \mathbf{m}_l^t$, with elements $M_{k,l}[m, n] = H^2[l, m] H^2[k, n]$ and observing

$$\mathbf{\Lambda} = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{M}_{k,l} \sigma_{\mathbf{E}_Y}^2[k, l]. \quad (11)$$

Here, $\mathbf{\Lambda}$ is the $N \times N$ array of error variances in the spatial domain. Thus the variance of the spatial-domain errors due to quantization with the 2-D DCT consists of a sum of the error basis images defined above, where each basis image is scaled by the error variance of its corresponding 2-D DCT coefficient. Figure 4 shows the error variance basis images for $N = 8$. Note that the error variance image will in general be spatially varying in m and n , as discussed previously for the 1-D case, and from (9) the error will in general be correlated.

Finally, from (5) each spatial-domain noise term is a linear combination of independently distributed random variables (64 of them, when using the 8×8 block DCT), allowing the spatial-domain quantization noise to be approximated as a 0-mean Gaussian random vector with autocovariance matrix $\mathbf{K}_{\mathbf{e}_z}$,

$$p(\mathbf{e}_z) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{K}_{\mathbf{e}_z}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \mathbf{e}_z^t \mathbf{K}_{\mathbf{e}_z}^{-1} \mathbf{e}_z \right\}. \quad (12)$$

(Note that the errors \mathbf{e}_y are assumed zero-mean here.)

²The authors do not make explicit use of (10) in their artifact reduction algorithm, but rather include it in an explanatory sense.

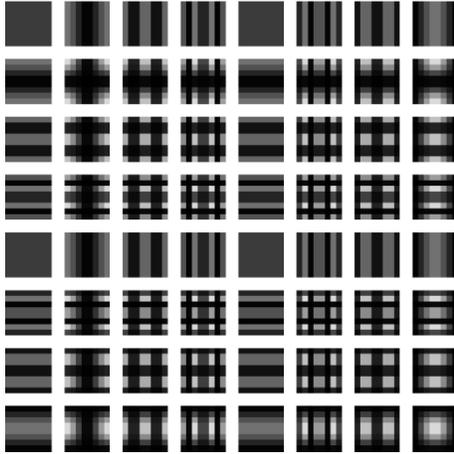


Fig. 4. Two-dimensional DCT error variance images for 8×8 image blocks. The basis images are placed with increasing horizontal frequency from left to right, and increasing vertical frequency from top to bottom. Basis image values are ordered with 0 at black, and increase as the image gets brighter.

As a final note for this section, it is important to recognize that although the spatial-domain error is modeled as Gaussian, the frequency-domain error is not: Indeed, the two errors e_z and e_y are related by the DCT; however, the Gaussian approximation for e_z arises due to the summation of the many individual error terms of e_y . The following section discusses the actual (non-Gaussian) statistical behavior of the frequency-domain error.

V. DCT-DOMAIN QUANTIZATION ERROR

In the previous section the autocovariance matrix of the spatial-domain quantization noise was derived, and depended on the frequency-domain quantization noise, which was assumed known. Here this missing quantity is provided. For the sake of brevity, only the two-dimensional case is discussed.

A simple model for frequency-domain quantization noise is based on the uniform distribution. If the original random variable representing the $(k, l)^{th}$ DCT coefficient is denoted as $Y[k, l]$, then it is quantized as $Y_q[k, l] = Q[Y[k, l]]$. If a realization $Y_q[k, l]$ is observed such that $Y_q[k, l]$ lies in the quantization interval $[q_i^{k,l}, q_{i+1}^{k,l})$, then the uniform model simply says that the quantization error $(Y_q[k, l] - Y[k, l])$ is distributed uniformly in the interval $[q_i^{k,l} - Y_q[k, l], q_{i+1}^{k,l} - Y_q[k, l])$. From an intuitive point of view, the uniform model says that an original frequency coefficient $Y[k, l]$ was equal to any of the values in this quantization range with equal probability. For the uniform model, the DCT-domain quantization error of the $(k, l)^{th}$ coefficient can be expressed

as

$$\sigma_{\mathbf{E}_Y}^2[k, l] = \frac{(q_{i+1}^{k,l} - q_i^{k,l})^2}{12}. \quad (13)$$

Using the uniform model makes analysis simple, for the variance of a uniform random variable is easy to calculate. Furthermore, the uniform model satisfies the conditions mentioned at the beginning of Section IV for the uncorrelatedness of \mathbf{E}_Y . However, although the uniform model makes analysis much simpler, it is not necessarily always appropriate. In many ways, the uniform model is similar to using the QCS constraints, and as a result can suffer from some of the same limitations as that approach.

Experimental simulations also suggest that using a uniform noise model can result in poor estimates of quantization noise. Rather than assuming uniform DCT-domain quantization noise, prior knowledge of the DCT coefficients can be incorporated in the model, which can yield results that are more accurate. By first considering an image model in the spatial domain, the desired result in the frequency domain will become available.

A model commonly used for the AC components of an image block \mathbf{Z} makes use of the separable autocorrelation function

$$E[Z[m, n] Z[i, j]] = \sigma_{\mathbf{Z}}^2 \rho_1^{|m-i|} \rho_2^{|n-j|}, \quad (14)$$

where the parameters ρ_1 and ρ_2 are the one-step correlations in the vertical and horizontal dimensions, $\sigma_{\mathbf{Z}}^2$ is the variance of the image block \mathbf{Z} , and the DC component of the block has been subtracted out, so that $E[\mathbf{Z}] = \mathbf{0}$. Using such a model, the variance of the $(k, l)^{th}$ DCT coefficient can be expressed as [16]

$$\sigma_{\mathbf{Y}}^2[k, l] = \sigma_{\mathbf{Z}}^2 [\mathbf{H}\mathbf{K}_1\mathbf{H}^t]_{k,k} [\mathbf{H}\mathbf{K}_2\mathbf{H}^t]_{l,l}, \quad (15)$$

where \mathbf{K}_x are defined as

$$\mathbf{K}_x = \begin{bmatrix} 1 & \rho_x & \rho_x^2 & \cdots & \rho_x^{N-1} \\ \rho_x & 1 & & & \\ \rho_x^2 & & \ddots & & \vdots \\ \vdots & & & & \rho_x \\ \rho_x^{N-1} & \cdots & \rho_x & 1 & \end{bmatrix}. \quad (16)$$

From (15) the variance, or energy, of each DCT coefficient is known provided that ρ_1 , ρ_2 , and σ_Z^2 are known. In practice, these parameters can be estimated from the observed data, or assumed to be known *a priori*.

The nature of the distributions of the DCT coefficients is of high importance for determining frequency-domain quantization noise. There is precedence for using a Laplacian model for DCT coefficients in images and video [20, 21]. The variance of a Laplacian random variable with parameter λ is $\sigma^2 = \frac{2}{\lambda^2}$, and the Laplacian parameter for the $(k, l)^{th}$ coefficient is found as

$$\lambda_{k,l} = \sqrt{\frac{2}{\sigma_Y^2[k, l]}}, \quad (17)$$

and the distribution of $Y[k, l]$ becomes

$$p_{Y[k,l]}(y) = \frac{\lambda_{k,l}}{2} \exp\{-\lambda_{k,l}|y|\}. \quad (18)$$

If the quantized DCT coefficient $Y_q[k, l] = y_q$ is observed to lie in the interval $[q_i^{k,l}, q_{i+1}^{k,l})$, then the distribution of $Y[k, l]$ given $Y_q[k, l] = y_q$ is zero outside the quantization interval and $\gamma p_{Y[k,l]}(y)$ inside the quantization interval, where γ is a constant that ensures the distribution integrates to unity. Thus,

$$\sigma_{E_Y}^2[k, l] = \gamma \int_{q_i^{k,l}}^{q_{i+1}^{k,l}} (y_q - y)^2 p_{Y[k,l]}(y) dy, \quad (19)$$

which can be determined in closed form via integration by parts.

Equations (13) and (19) give two possible methods for determining the frequency-domain quantization noise. The method of (13) uses a uniform distribution for the DCT coefficients, which is equivalent to assuming equal probability for all coefficients that lie inside the quantization limits. Alternatively, the equation in (19) was derived based on a prior model for the image.

Experimental simulations have suggested that by themselves, neither (13) nor (19) accurately predict the spatial-domain quantization noise. For quantized DCT coefficients that are observed to be zero, the Laplacian model indeed works well. However, when non-zero quantized DCT coefficients are observed, the Laplacian model seems to be inappropriate and the uniform model works better.³ Using the uniform model for non-zero coefficients can be viewed as approximating the tails of the Laplacian with a uniform distribution, rather than the exponentially decaying

³The work reported by Lakhani [22] also seems to support this: The author reports only marginal improvements from an algorithm that decodes non-zero DCT coefficients based on a Laplacian model.

tail. Using different distributions depending on whether or not the observed quantized coefficient is zero essentially says that the expected energy of the DCT coefficients is concentrated near zero, and when known to be away from zero (i.e., $Y_q[k, l] \neq 0$), the energy is expected to be more evenly distributed. Such a distribution also satisfies the conditions mentioned at the beginning of Section IV that ensure the errors \mathbf{E}_Y are uncorrelated, whereas using only the Laplacian model does not necessarily do so.

VI. INTERPRETATION AND IMPLICATIONS

Consider the 1-D error variance basis vectors of Fig. 3. If the DCT quantization errors were equal for each frequency, then the basis sum of (8) would result in equal variance at each spatial location. However, in general the DCT quantization errors will not be equal at each frequency. For example, for relatively smooth signals more signal energy will be contained in the low-frequency coefficients, with decreasing energy as the frequency increases. The high-frequency coefficients are quantized to zero, but since they have little energy, they contribute very little quantization noise. However, the low-frequency coefficients contribute considerable quantization noise. Thus, the sum of (8) will have higher weights for the low-frequency components, and a quick consultation with Fig. 3 suggests that this will result in higher error variance for locations near the boundaries. This simple argument accounts for the presence of blocking artifacts, and has been pointed out before by Choi and Kim [19]. Note also from Fig. 3 that errors in the high-frequency components contribute little to the error right at the block boundaries.

On the other hand, consider signals that do contain significant high-frequency content, such as textured regions of an image. For these cases the situation described above is reversed, and examples can easily be constructed for which pixels near the boundaries may actually have less error variance than pixels near the middle.

The 2-D case is a ready extension of the case in one dimension, with similar arguments for the existence or absence of blocking artifacts as in one dimension. While not as obvious as the 1-D case, the error variance basis images of Fig. 4 exhibit the same properties as the basis vectors in Fig. 3: High error at block boundaries for low-frequency components, and low error at block boundaries for high-frequency components. From (15), positive values of the correlation parameters ρ_1 and ρ_2 lead to concentrations of signal energy at low frequencies, which leads to increased error near block boundaries. Similarly, negative values of the correlation parameters

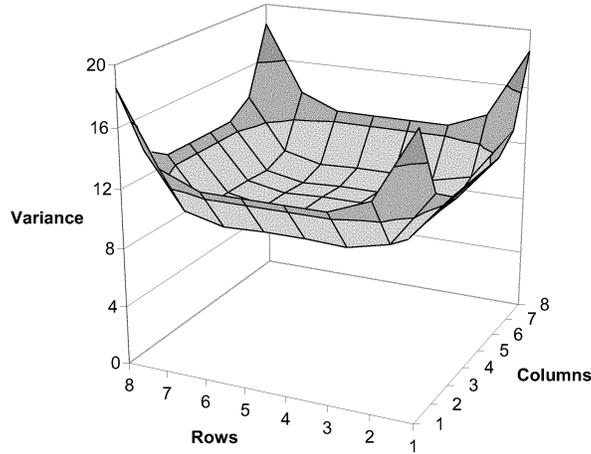


Fig. 5. Predictions for 2-D error variance as a function of spatial location using $\rho_1 = \rho_2 = 0.8$, $\sigma_Z^2 = 25$, $QP = 12$, and assuming non-zero DCT coefficients only in locations $(0,0)$, $(0,1)$, and $(1,0)$.

lead to concentrations of signal energy at the high frequencies, leading to quantization errors that are lower at block boundaries than in block interiors. As an example, consider the plot in Fig. 5 that shows the 2-D spatial-domain error variance predicted by the proposed model when $\rho_1 = \rho_2 = 0.8$, $\sigma_Z^2 = 25$, $QP = 12$, and only the three lowest-frequency components are quantized to non-zero values (H.263-style quantization is used). These parameters correspond to a relatively smooth image, and the results of the figure confirm the intuitive belief that the error variance is large at block boundaries and block corners, as was true for the example in Fig. 2. (Note however that one cannot expect the two figures to coincide perfectly, because the results in Fig. 2 were averaged over all blocks in a sequence, and the prediction in Fig. 5 is for a particular block with specific characteristics.)

This analysis gives theoretic justification for many of the ad hoc post processing algorithms that attempt to remove blocking artifacts. Many of these algorithms perform selective smoothing at block boundaries, where more smoothing is applied in regions that are deemed “smooth,” and less or no smoothing in regions that are “not so smooth.” Filtering block edges like this makes perfect sense in light of the discussion on the presence or absence of blocking artifacts.

Extensive tests have been performed for DCT compression noise in one dimension, and the results support the discussions of this paper—quantization noise in the spatial domain is undoubtedly spatially varying and correlated, with a strong dependence on the quantization intervals and on the original signal statistics. Figure 6 shows results obtained experimentally for the

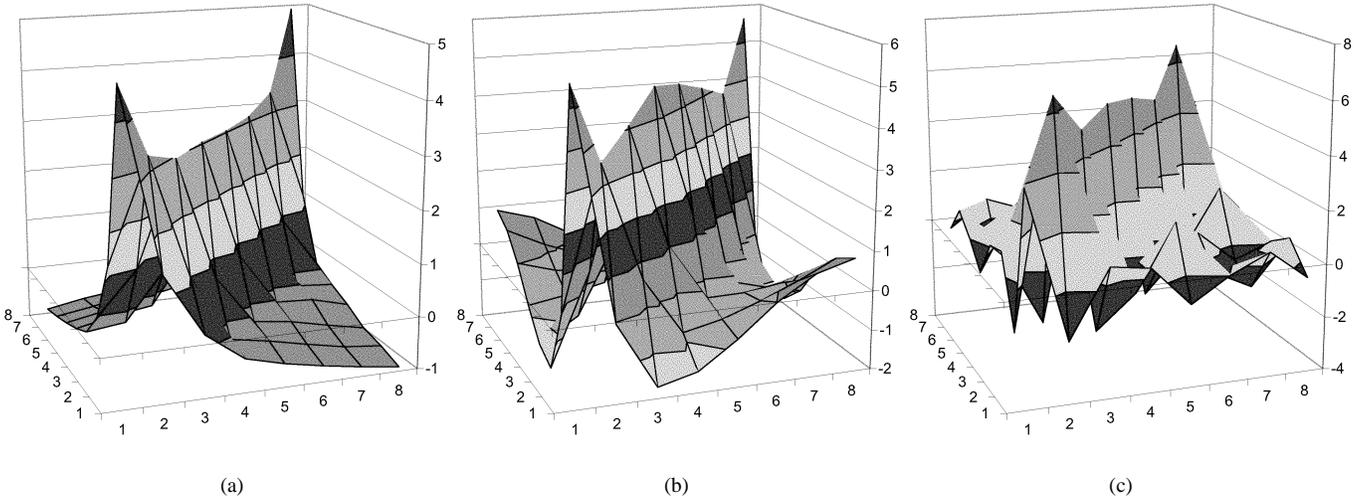


Fig. 6. Experimentally observed quantization error covariance matrices for 1-D image scan lines of length 8 in the 176×144 *foreman* sequence. These results have been accumulated and averaged for all vectors with non-zero entries in the lowest (a) one, (b) two, and (c) five DCT frequency coefficients. Each of these three results was obtained based on over 60,000 observations.

1-D case, where a number of length-8 image scan lines from the 176×144 *foreman* sequence have been compressed and the error for vectors with various patterns of zero/non-zero quantized DCT coefficients was recorded. Quantization intervals of length 6.0 were used for the DCT coefficients, with the exception of values in the range $(-6.0, 6.0)$, which were quantized to 0. The plot in Fig. 6(b), for example, shows the covariance of the quantization error for all observations that had non-zero quantized values only in the two lowest-frequency DCT coefficients. The covariance was computed using the standard estimation formula,

$$\hat{\mathbf{K}} = \frac{1}{W} \sum_{i=0}^{W-1} \mathbf{e}_i \mathbf{e}_i^t, \quad (20)$$

where for the case of Fig. 6(b) \mathbf{e}_i , $i = 0, \dots, W - 1$ are the errors of all the observations that had non-zero DCT coefficients in the two lowest-frequency components, and zero DCT coefficients otherwise. Note that by computing the covariance like this the statistics of the noise are assumed constant for a particular pattern of quantized zero/non-zero DCT coefficients, and that the statistics may vary with different patterns.

There are several observations to make about the results in Fig. 6. One important note concerns the diagonal elements, for these are the values corresponding to the variance of the quantization error at each of the eight spatial locations. Both (a) and (b) of the figure correspond

to relatively smooth vectors, since only the one or two lowest frequencies were observed to be non-zero. For both of these cases, it is obvious that the error variance is significantly higher at the vector boundaries, supporting the arguments made earlier in this section. Another observation from these two plots is that they both have significant values in the non-diagonal elements, meaning that the error is most definitely not independent.

The plot in part (c) of Fig. 6 is also instructive to examine. Plot (c) has results for vectors that contained considerably more high-frequency energy than in (a) and (b), and its differences with these two are apparent. Details of plot (c) are difficult to see due to the varying nature of its data, and therefore these data are also explicitly tabulated in Table I. From the table, it is quickly apparent that the variances of the error near the vector boundaries are considerably less than those in the vector interior. As a final note, the quantization error is also clearly not independent.

Similar experiments were conducted for the 2-D case. The same conclusions, that the spatial-domain errors due to quantization are correlated with a spatially-varying variance, can be drawn. This is expected since the 2-D case is simply a separable extension of the case in one dimension. Covariance matrices in two dimensions are not presented because the results are difficult to visualize due to the necessity of using the stacked-column notation, and in the end they do not provide significantly more information than that of the 1-D results already presented.

VII. APPLICATION: ARTIFACT REMOVAL

In this section, the noise model introduced in this paper is incorporated in a post-processing scheme to remove compression artifacts. The probabilistic description of DCT quantization noise derived in previous sections is well suited to a stochastic image restoration formulation, and that is the path followed here. In this formulation, the quantization noise model becomes a *likelihood* term that ensures that the final image estimate agrees with the observed data. The restoration method is Bayesian, in the sense that the formulation introduces an *a priori* term to smooth out compression artifacts.

For other methods of artifact reduction in compressed images, the reader is directed to many of the references cited previously in different contexts, in particular [2, 6–12, 19, 22]

TABLE I
ENTRIES FOR THE COVARIANCE MATRIX IN FIG. 6(C).

Estimated Covariance Matrix							
3.87	-1.70	0.82	0.60	-0.89	0.03	0.71	-0.41
-1.70	6.58	-2.03	-0.78	1.69	-0.34	-1.05	0.65
0.82	-2.03	4.96	-0.81	-0.25	0.72	-0.48	0.11
0.60	-0.78	-0.81	5.51	-1.95	-0.41	1.92	-1.01
-0.89	1.69	-0.25	-1.95	5.32	-0.65	-0.86	0.62
0.03	-0.34	0.72	-0.41	-0.65	4.86	-2.08	0.86
0.71	-1.05	-0.48	1.92	-0.86	-2.08	6.67	-1.79
-0.41	0.65	0.11	-1.01	0.62	0.86	-1.79	3.86

A. MAP Estimate

A maximum *a posteriori* (MAP) criterion is used for this work, which seeks the solution

$$\hat{\mathbf{z}} = \arg \max_{\mathbf{z}} p(\mathbf{z} | \mathbf{z}_q), \quad (21)$$

$$= \arg \max_{\mathbf{z}} p(\mathbf{z}) p(\mathbf{z}_q | \mathbf{z}). \quad (22)$$

The latter probability of this last equation is the likelihood term, whereas the former is the *a priori* term for the frame under consideration. The likelihood term is formed by considering (12): since $\mathbf{z}_q = \mathbf{z} + \mathbf{e}_z$, $\mathbf{z}_q | \mathbf{z}$ is a Gaussian random variable with mean \mathbf{z} and autocovariance $\mathbf{K}_{\mathbf{e}_z}$. Strictly speaking, \mathbf{z}_q is a deterministic function of \mathbf{z} ; however, as mentioned at the beginning of Section III it is more productive to treat the errors as random quantities, where the random quantities are described according to the error analysis of previous sections. For the likelihood term, the uniform frequency-domain noise model described by (13) is chosen instead of the Laplacian model of (19). This avoids the added complexity and difficulty of accurately estimating the parameters ρ_1 , ρ_2 , and σ_Z^2 on a block-by-block basis. As will be seen, using the uniform model provides good post-processing results while demonstrating the effectiveness of the proposed quantization noise model.

Note that the vectors in (22) correspond to entire images, whereas previously they only denoted individual image blocks. Thus $p(\mathbf{z})$ will be a prior probability for the entire image, and $p(\mathbf{z}_q | \mathbf{z})$ is constructed on a block-by-block basis according to the noise model introduced in previous sections. Note that $\mathbf{K}_{\mathbf{e}_z}$ in this case will have a block-diagonal structure, with blocks equal

to the error covariances for each individual image block within the image; $\mathbf{K}_{\mathbf{e}_y}$ is constructed analogously, and consists of the error variances of all DCT coefficients in the image. The notation \mathcal{D} and \mathcal{D}^t represents the block DCT and IDCT, which are block-diagonal matrices that perform the DCT and IDCT on each individual block of an image vector. Each matrix within the block-diagonal \mathcal{D} is the 2-D DCT matrix \mathcal{H} ; the block IDCT consists of repetitions of \mathcal{H}^t along its diagonal.

The prior model will be based on the Huber-Markov Random Field (HMRF), which has been used extensively in image and video processing; for examples, see the references [6, 23, 24]. The Huber function is a convex function that has edge-preserving properties relative to a simple quadratic. Details of the HMRF can be found in the above references, and is given here without excessive discussion,

$$p(\mathbf{z}) = \frac{1}{G} \exp \left\{ -\lambda \sum_{c \in \mathcal{C}} \rho_T(\mathbf{d}_c^t \mathbf{z}) \right\}, \quad (23)$$

where G is a normalizing constant, λ is a regularization parameter, and the Huber function $\rho_T(\cdot)$ is defined as

$$\rho_T(u) = \begin{cases} u^2, & |u| \leq T, \\ T^2 + 2T(|u| - T), & |u| > T. \end{cases} \quad (24)$$

The c of (23) are local groups of pixels called cliques, and \mathcal{C} is the set of all such cliques, which depends on the neighborhood structure of the MRF. Here, the vectors \mathbf{d}_c are chosen to extract the differences between a pixel and its neighbors, such that (23) simplifies to

$$p(\mathbf{z}) = \frac{1}{G} \exp \left\{ -\lambda \sum_{n=0}^{M-1} \sum_{m \in \mathcal{N}_n} \rho_T(z[n] - z[m]) \right\}, \quad (25)$$

where \mathcal{N}_n is the index set of neighbors for the n^{th} pixel, and M is the number of pixels in the image. The inner summation in (25) is over each pixel in the neighborhood of n . A neighborhood consisting of the eight nearest neighbors of a pixel is used in this example.

Maximizing the term in (22) reduces to

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} [\lambda r(\mathbf{z}) + s(\mathbf{z})], \quad (26)$$

where

$$r(\mathbf{z}) = \sum_{c \in \mathcal{C}} \rho_T(\mathbf{d}_c^t \mathbf{z}), \quad (27)$$

$$s(\mathbf{z}) = \frac{1}{2} (\mathbf{z} - \mathbf{z}_q)^t \mathbf{K}_{\mathbf{e}_z}^{-1} (\mathbf{z} - \mathbf{z}_q). \quad (28)$$

The optimization problem in (26) is solved using a gradient descent algorithm [25]. The gradient of the objective function is

$$\mathbf{g}(\mathbf{z}) = \lambda \nabla r(\mathbf{z}) + \nabla s(\mathbf{z}), \quad (29)$$

where the individual gradients of the two terms with respect to the vector \mathbf{z} are, after substituting for $\mathbf{K}_{\mathbf{e}_z}$,

$$\nabla r(\mathbf{z}) = \sum_{c \in \mathcal{C}} \mathbf{d}_c \rho'_T(\mathbf{d}_c^t \mathbf{z}), \quad (30)$$

$$\nabla s(\mathbf{z}) = \mathcal{D}^t \mathbf{K}_{\mathbf{e}_y}^{-1} \mathcal{D}(\mathbf{z} - \mathbf{z}_q). \quad (31)$$

Both of these terms have importance, and warrant further examination. The gradient term in (30) has larger values for large differences in $\mathbf{d}_c^t \mathbf{z}$, thus encouraging smoothness. Note, however, that the first derivative of the Huber function has a maximum magnitude of $2T$, which effectively prevents the gradient from becoming too large, and thus prevents excessive smoothing of image edges in the final result.

The effect of Equation (31) is such that DCT frequency components that have larger variance (due to the larger quantization intervals in (13)) do not affect the gradient as much as DCT frequency components with lower variance (due to smaller quantization intervals in (13)). Thus the proposed noise model automatically adjusts for different quantizers—if the quantization parameters decrease for certain regions of the frame (due to, for example, some sort of region-of-interest coding [26]), the model accounts for this by having lower quantization noise for those regions. Similarly, if the quantization parameters change from frame to frame, the quantization noise for each frame follows accordingly. This is a much more realistic model than assuming IID noise throughout each frame.

Denoting the estimate of \mathbf{z} at iteration w as $\mathbf{z}^{(w)}$, the gradient descent algorithm forms the new estimate as

$$\mathbf{z}^{(w+1)} = \mathbf{z}^{(w)} - \alpha^{(w)} \mathbf{g}(\mathbf{z}^{(w)}), \quad (32)$$

where $\alpha^{(w)}$ is a step size that ideally reduces the objective function by as much as possible. The step size is determined by a simple one-dimensional search algorithm that finds the best $\alpha^{(w)}$ in a pre-defined search range, where the search is performed at a pre-determined resolution. The iterations of (32) continue until improvements in the objective function fall below a small threshold.

TABLE II
 COMPARISON PSNR RESULTS FOR POST-PROCESSING SEVERAL SEQUENCES COMPRESSED
 ACCORDING TO THE H.263 STANDARD AT 30 FPS USING ONLY I-FRAMES.

Sequence	Size	Frames	QP ¹	λ	Δ PSNR, dB ²		
					proposed ³	O'Rourke ³	Annex J
<i>foreman</i>	176 × 144	0–275	16	0.00075	+1.12	+0.33	+0.38
<i>foreman</i>	176 × 144	0–275	12	0.00075	+1.01	+0.03	+0.34
<i>claire</i>	320 × 272	0–167	16	0.00075	+0.83	+0.05	+0.56
<i>claire</i>	320 × 272	0–167	12	0.00075	+0.60	−0.24	+0.51
<i>missa</i>	352 × 288	0–149	18	0.00075	+0.60	+0.28	+0.56
<i>missa</i>	352 × 288	0–149	14	0.00075	+0.38	+0.11	+0.46

¹QP is the quantization parameter defined in the H.263 standard

² Δ PSNR is the average change in PSNR per frame for the sequence

³The Huber parameter is $T = 10.0$

Using the algorithm discussed in this section, it is possible to get image estimates that do not satisfy the QCS. If such situations are of concern to one implementing the algorithm, it is a simple matter to include a projection to the QCS of each image estimate during the iterations of (32).

It should be noted that significant computational improvements over this simple approach can be made [27]; however, computational efficiency is not a focus of this paper.

B. Experimental Results

The post-processing algorithm described above has been implemented in an H.263 video decoder. Results are presented only for intra-coded frames (I-frames) so that DCT quantization noise for images is not confused with quantization noise for motion-compensated residual images, because the statistics of the two are quite different. All sequences discussed here have been compressed to 30 fps with a constant QP, and the resulting PSNR values are the average of each frame's PSNR. Table II shows results for three sequences compressed with various



Fig. 7. Example post-processing results for *foreman* intra frame. Part (a) shows the unprocessed frame 63, QP=16, and the remaining frames are processed with (b) proposed algorithm, $\lambda = 0.00075$; (c) Annex J; (d) O'Rourke.

QP's. For comparison purposes, Table II also shows results of post-processing with two other algorithms—the MAP algorithm of O'Rourke [6], and the Annex J algorithm of the H.263 standard [15] implemented as a post-filter (as opposed to implementation within the coding loop as specified in the standard). The Annex J filter is a simple de-blocking filter that filters across block boundaries; when applied to a vertical block edge, it modifies two pixels to the left of the border and two pixels to the right of the border for each row. The strength of the filter is dependent on the severity of the quantization parameters used for the blocks being processed. Horizontal block edges are filtered analogously.



Fig. 8. Example post-processing results for *claire* intra frame. Part (a) shows the unprocessed frame 50, QP=16, and the remaining frames are processed with (b) proposed algorithm, $\lambda = 0.00075$; (c) Annex J; (d) O'Rourke.

Several conclusions can be drawn from the table, not least of which are the average PSNR improvements—in all cases, significant improvements have been made in this objective measure, and the proposed algorithm generally (but not always) provides superior PSNR results relative to the other two algorithms. Another aspect of the table that draws attention is the column of values for λ , the regularization parameter—the same value has been used in each case. This has been done deliberately to show an important characteristic of the proposed quantization noise model: Since the noise model is directly dependent on the quantization intervals, it automatically

adjusts the amount of expected noise according to the received quantization information. This means that once a particular λ has been found that is appropriate for a certain class of images, that same λ can be used in reconstructing those images for a wide range of quantization severity. This does not mean that this λ provides the optimal PSNR for the experiments presented in the table; indeed, slightly varying λ for each case reported in Table II would have resulted in further small PSNR improvements (but little noticeable visual difference). Note that λ is simply a parameter of the HMRF model that controls how smooth we assume the image is *a priori*. The overall amount of smoothing to be done depends on both λ and the quantization severity in the image. If one holds λ constant and varies the quantization severity, the apparent amount of smoothing will vary according to the quantization.

An important implication of this can be seen by considering what happens when the QP changes values within a frame, as opposed to between frames. Quantization levels can change within a frame when, for example, one performs region-of-interest coding—the encoder may decide to compress some “regions of interest” with a much lower QP than the rest of the frame in order to have improved quality in these regions at the cost of quality elsewhere. In post-processing frames compressed like this, more smoothing may be necessary in the regions compressed with coarser quantization levels than in the regions of interest. Algorithms that have IID noise models would have to vary a regularization parameter across the frame to allow more or less smoothing, which seems counterintuitive since the reason more smoothing is necessary is because there is more quantization noise, not because the original image is actually smoother there.

While Table II showed objective PSNR results for the proposed post-processing algorithm, it is easy to argue that PSNR, while being “objective,” is not necessarily the best metric with which to evaluate an image restoration scheme—the bottom line in any image restoration scheme is the visual quality of the results. Figures 7 and 8 show close-ups of some example frames from the experiments reported in Table II, and it is readily evident that the algorithm effectively removes the blocking artifacts from the received images. Furthermore, over-smoothing is not apparent in these reconstructions. Blocking artifacts have been successfully removed in both the Annex J images and O’Rourke’s images, but ringing artifacts are still present for Annex J, while slight oversmoothing is noticeable in O’Rourke’s algorithm.

The quantization noise model introduced in this paper can be applied in a wide range of restoration schemes of DCT-compressed images and video. Other examples include temporal resolution enhancement of compressed video (or temporal interpolation, i.e., increasing the frame rate) [28, 29], and restoration of inter-compressed video frames (e.g., H.263 or MPEG P-Frames) [29].

VIII. CONCLUSION

This paper has provided a statistical description of the spatial-domain quantization noise in images compressed using the DCT. Both the variance of the quantization error at individual pixels, as well as the correlations between them, have been considered. It was demonstrated that the spatial-domain quantization noise is in general correlated and spatially varying. The presence of blocking artifacts became easily explicable in this context, as was their absence. It was shown that in some cases one can actually expect pixel values near block boundaries to be more accurate than those in the block interior.

An understanding of compression noise is crucial if one expects to be able to compensate appropriately for such noise, making the concepts discussed in this paper valuable for those in the compressed image and video post-processing fields. In particular, the statistical descriptors of compression noise derived in this paper are readily applicable in a probabilistic image restoration formulation, and this paper provided one such example that showed both quantitatively and qualitatively the utility of the proposed noise model.

REFERENCES

- [1] M. A. Robertson and R. L. Stevenson. DCT quantization noise in compressed images. In *International Conf. on Image Processing*, volume 1, pages 185–188, Oct. 7–12 2001.
- [2] M. A. Robertson and R. L. Stevenson. Restoration of compressed video using temporal information. In *SPIE Visual Communications and Image Processing*, volume 4310, pages 21–29, Jan. 24–26 2001.
- [3] A. J. Ahumada and R. Horng. De-blocking DCT compressed images. In *SPIE Human Vision, Visual Processing, and Digital Display*, volume 2179, pages 109–116, 1994.
- [4] H. A. Peterson, A. J. Ahumada, and A. B. Watson. The visibility of DCT quantization noise. In *J. Marreale, ed., Digest of Technical Papers, Society for Information Display, Playa del Rey, CA*, volume 24, pages 942–945, 1993.
- [5] B. Widrow, I. Kollár, and M.-C. Liu. Statistical theory of quantization. *IEEE Trans. Instrumentation and Measurement*, 45(2):353–361, Apr. 1996.
- [6] T. P. O’Rourke and R. L. Stevenson. Improved image decompression for reduced transform coding artifacts. *IEEE Trans. Circuits and Systems for Video Technology*, 5(6):490–499, Dec. 1995.

- [7] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos. Projection-based spatially adaptive reconstruction of block-transform compressed images. *IEEE Trans. Image Processing*, 4(7):896–908, July 1995.
- [8] S. H. Park and D. S. Kim. Theory of projection onto the narrow quantization constraint set and its applications. *IEEE Trans. Image Processing*, 8(10):1361–1373, Oct. 1999.
- [9] T. Meier, K. N. Ngan, and G. Crebbin. Reduction of blocking artifacts in image and video coding. *IEEE Trans. Circuits and Systems for Video Technology*, 9(3):490–500, Apr. 1999.
- [10] J. Mateos, A. K. Katsaggelos, and R. Molina. A Bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts. *IEEE Trans. Image Processing*, 9(7):1200–1215, July 2000.
- [11] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos. Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images. *IEEE Trans. Circuits and Systems for Video Technology*, 3(6):421–432, Dec. 1993.
- [12] J. Yang, H. Choi, and T. Kim. Noise estimation for blocking artifacts reduction in DCT coded images. *IEEE Trans. Circuits and Systems for Video Technology*, 10(7):1116–1120, Oct. 2000.
- [13] H. A. Peterson, A. J. Ahumada, and A. B. Watson. The visibility of DCT quantization noise: Spatial frequency summation. In *Proc. Data Compression Conference*, page 530. IEEE Computer Press, March 26–31 1994.
- [14] G. S. Yovanof and S. Liu. Statistical analysis of the DCT coefficients and their quantization error. In *Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 601–605, Nov. 3–6 1996.
- [15] ITU-T Recommendation H.263. *Video coding for low bit rate communication*, 1998.
- [16] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1989.
- [17] L. Karray, P. Duhamel, and O. Rioul. Image coding with an L^∞ norm and confidence interval criteria. *IEEE Trans. Image Processing*, 7(6):621–631, May 1998.
- [18] A. B. Sripad and D. L. Snyder. A necessary and sufficient condition for quantization errors to be uniform and white. *IEEE Trans. Acoust., Speech, and Signal Processing*, ASSP-25(5):442–448, Oct. 1977.
- [19] H. Choi and T. Kim. Blocking-artifact reduction in block-coded images using wavelet-based subband decomposition. *IEEE Trans. Circuits and Systems for Video Technology*, 10(5):801–805, Aug. 2000.
- [20] R. C. Reininger and J. D. Gibson. Distributions of the two-dimensional DCT coefficients for images. *IEEE Trans. on Communications*, COM-31(6):835–839, June 1983.
- [21] E. Y. Lam and J. W. Goodman. A mathematical analysis of the DCT coefficient distributions for images. *IEEE Trans. Image Processing*, 9(10):1661–1666, Oct. 2000.
- [22] G. Lakhani. Distribution-based restoration of DCT coefficients. *IEEE Trans. Circuits and Systems for Video Technology*, 10(5):819–823, Aug. 2000.
- [23] R. R. Schultz and R. L. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Trans. Image Processing*, 5(6):996–1011, June 1996.
- [24] R. R. Schultz and R. L. Stevenson. A Bayesian approach to image expansion for improved definition. *IEEE Trans. Image Processing*, 3(3):233–242, May 1994.
- [25] S. L. S. Jacoby, J. S. Kowalik, and J. T. Pizzo. *Iterative Methods for Nonlinear Optimization Problems*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1972.
- [26] L. L. Yang and M. A. Robertson. Multiple-face tracking system for general region-of-interest video coding. In *International Conf. on Image Processing*, volume 1, pages 347–350, Sep. 10–13 2000.
- [27] M. A. Robertson and R. L. Stevenson. Reduced-complexity iterative post-filtering of video. *IEEE Trans. Circuits and Systems for Video Technology*, 11(10):1121–1127, Oct. 2001.
- [28] M. A. Robertson and R. L. Stevenson. Temporal resolution enhancement in compressed video sequences. *EURASIP*

Journal on Applied Signal Processing: Special Issue on Nonlinear Signal and Image Processing, 2001(4):230–238, Dec. 2001.

- [29] M. A. Robertson. *High-Quality Reconstruction of Digital Images and Video from Imperfect Observations*. Ph.D. dissertation, University of Notre Dame, Apr. 2001.



Mark A. Robertson received the B.S. in Electrical Engineering from Tri-State University, Angola, IN, in 1996, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Notre Dame, Notre Dame, IN, in 1998 and 2001. While at Notre Dame, he was supported by the Arthur J. Schmitt Fellowship, the Intel Corporation, and the Indiana Space Grant Consortium. His current research interests include image and video enhancement, image and video communication, and aerial surveillance applications.



Robert L. Stevenson was born on December 20, 1963, in Ridley Park, Pennsylvania. He received the B.E.E. degree (summa cum laude) from the University of Delaware in 1986, and the Ph.D. in Electrical Engineering from Purdue University in 1990. While at Purdue he was supported by graduate fellowships from the National Science Foundation, DuPont Corporation, Phi Kappa Phi, and Purdue University. He joined the faculty of the Department of Electrical Engineering at the University of Notre Dame in 1990, where he is currently an Professor. Dr. Stevenson also holds a co-current appointment as a Professor in the Department of Computer Science and Engineering at the University of Notre Dame.

His research interests include image/video processing, image/video compression, robust image/video communication systems, multimedia systems, ill-posed problems in computational vision, and computational issues in image processing.

Dr. Stevenson has served as an Associate Editor for the IEEE Transactions on Image Processing, the IEEE Transactions on Circuits and Systems for Video Technology, and the Journal of Electronic Imaging.

Dr. Stevenson is a member of the IEEE, SPIE, Eta Kappa Nu, Tau Beta Pi, and Phi Kappa Phi.

LIST OF FIGURES

1	Different images that each lie in the same quantization constraint set of a P-frame of the <i>foreman</i> sequence: (a) the original image; (b) compressed image with standard de-compression; (c) compressed image that has been smoothed subject to quantization constraints; and (d) compressed image that has been sharpened subject to quantization constraints.	7
2	Average quantization error variance for each block of the 320×272 <i>claire</i> sequence compressed using I-frames only, with H.263 quantization parameter 12.	9
3	One-dimensional basis vectors of the spatial-domain quantization error variance for sequences of length 8. The horizontal axis of each plot corresponds to spatial location n and the frequencies start with $k = 0$ for the top left and continue to $k = 7$ for the bottom right.	10
4	Two-dimensional DCT error variance images for 8×8 image blocks. The basis images are placed with increasing horizontal frequency from left to right, and increasing vertical frequency from top to bottom. Basis image values are ordered with 0 at black, and increase as the image gets brighter.	12
5	Predictions for 2-D error variance as a function of spatial location using $\rho_1 = \rho_2 = 0.8$, $\sigma_Z^2 = 25$, QP = 12, and assuming non-zero DCT coefficients only in locations (0,0), (0,1), and (1,0).	16
6	Experimentally observed quantization error covariance matrices for 1-D image scan lines of length 8 in the 176×144 <i>foreman</i> sequence. These results have been accumulated and averaged for all vectors with non-zero entries in the lowest (a) one, (b) two, and (c) five DCT frequency coefficients. Each of these three results was obtained based on over 60,000 observations.	17
7	Example post-processing results for <i>foreman</i> intra frame. Part (a) shows the unprocessed frame 63, QP=16, and the remaining frames are processed with (b) proposed algorithm, $\lambda = 0.00075$; (c) Annex J; (d) O'Rourke.	23
8	Example post-processing results for <i>claire</i> intra frame. Part (a) shows the unprocessed frame 50, QP=16, and the remaining frames are processed with (b) proposed algorithm, $\lambda = 0.00075$; (c) Annex J; (d) O'Rourke.	24

LIST OF TABLES

I	Entries for the covariance matrix in Fig. 6(c).	19
II	Comparison PSNR results for post-processing several sequences compressed according to the H.263 standard at 30 fps using only I-frames.	22

LIST OF FOOTNOTES

Manuscript received —.

Robert L. Stevenson is with the Laboratory for Image and Signal Analysis (LISA) in the Electrical Engineering Department at the University of Notre Dame. Mark A. Robertson was with the LISA, and is currently with the Air Force Research Laboratory.

Portions of this work appear in [1, 2].

This work was supported in part by Indiana's 21st Century Fund and by Sun Microsystems.

¹Equivalently, a regularization parameter can be varied to control the amount of smoothing to perform, where the regularization parameter affects the *a priori* smoothness assumed for the image.

²The authors do not make explicit use of (10) in their artifact reduction algorithm, but rather include it in an explanatory sense.

³The work reported by Lakhani [22] also seems to support this: The author reports only marginal improvements from an algorithm that decodes non-zero DCT coefficients based on a Laplacian model.