

# Recompression of JPEG Images by Requantization

Heinz H. Bauschke, *Member, IEEE*, Christopher H. Hamilton, *Student Member, IEEE*,  
Mason S. Macklem, *Student Member, IEEE*, Justin S. McMichael, and Nicholas R. Swart

**Abstract**—In this paper, we report a novel heuristic for requantizing JPEG images. The resulting images are generally smaller and often have improved perceptual image quality over a “blind” requantization approach, that is, one that does not consider the properties of the quantization matrices. The heuristic is supported by a detailed mathematical treatment which incorporates the well-known Laplacian distribution of the AC discrete cosine transform (DCT) coefficients with an analysis of the error introduced by requantization. We note that the technique is applicable to any image compression method which employs discrete cosine transforms and quantization.

**Index Terms**—Compression, JPEG image format, quantization, recompression, requantization.

## I. INTRODUCTION

THE JPEG image compression standard [1]–[5] is employed in a large number of image-intensive applications. In some of these applications, requantization is required when the amount of compression needed is unknown in advance. The Internet provides a good example: with the proliferation of different clients and connection speeds, it is often not possible to know *a priori* what a suitable compression level is. (This has led to the invention of, for example, transformational proxies, which transform web images on the fly based on real-time measurements of client properties and connection speed.) Ideally, one would like to always work from the original image when requantizing since JPEG compression is lossy. However, the original image is not always available: it may no longer exist or it may be too difficult to retrieve in real time.

Requantizing an already quantized image can lead to seemingly unpredictable behavior and unwanted artifacts. The following experiment can be reproduced using any photo-realistic image: consider the JPEG image shown in Fig. 1 which was obtained by compressing the test image Lena to quality level Q75, where  $Q_n$  denotes the  $n$ th quality setting suggested by the Independent JPEG Group (IJG) [5]. Requantizing the Q75 image to Q50, in other words not employing the original, gives the re-



Fig. 1. JPEG Q75 image Lena – 75.



Fig. 2. JPEG image Lena – 75 – 50, requantized from Lena – 75 to Q50.

Manuscript received August 23, 2001; revised January 20, 2003. This work was supported by Packeteer (R), Cupertino, CA. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Nasir Memon.

H. H. Bauschke is with the Department of Mathematics and Statistics, University of Guelph, Guelph, ON N1G 2W1, Canada (e-mail: hbauschk@uoguelph.ca).

C. H. Hamilton and M. S. Macklem are with the Department of Mathematics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada (e-mail: chamilt@cecm.sfu.ca; mmacklem@cecm.sfu.ca).

J. S. McMichael is with the Itiva Development Corporation, Kelowna, BC V1Y 7T2, Canada (e-mail: justin@itiva.com).

N. R. Swart is with the Analog Devices, Wilmington, MA 01887-3462 USA (e-mail: nicholas.swart@analog.com).

Digital Object Identifier 10.1109/TIP.2003.812375

sulting image in Fig. 2, which is 19% smaller, but has considerable visual artifacts: the Q50 version looks “grainier” than the Q75 image. Contrary to intuition, if the Q75 image is requantized to Q48, the grainy artifacts disappear, as seen in Fig. 3.



Fig. 3. JPEG image Lena – 75 – 48, requantized from Lena – 75 to Q48.

The Q48 image is perceptually much closer to the Q75 image than the Q50, yet is 37% smaller. In other words, this experiment shows that the IJG quality rating scale is not perceptually monotone. Similar phenomena were observed and discussed by Chan in [6] and [7].

In this paper we present a new heuristic algorithm that requantizes any JPEG image to one whose quantization matrices approximate fixed, user-specified quantization matrices so that the artifacts observed in the previous example are avoided. Our approach has application not only to the JPEG image compression standard, but to any method which employs a DCT-based transform step, such as the popular MPEG 1, MPEG 2 and MPEG 4 video coding algorithms [8, Sections 6.1, 6.4, 6.5].

The paper is organized as follows. The Laplacian distribution of the AC DCT coefficients is discussed in Section II. In Section III, the effect of requantizing an integer in two steps is studied. The error introduced by dequantizing a Laplacian distribution twice is analyzed in Section IV and applied to JPEG images in Section V. Section VI contains a detailed description of our new heuristic algorithm.

## II. DISTRIBUTION OF JPEG DCT COEFFICIENTS

### A. Main Steps of JPEG Image Compression

- 1) The image is separated into three color components.
- 2) Each component is partitioned into nonoverlapping  $8 \times 8$  blocks.
- 3) Each block is transformed using the two-dimensional Discrete Cosine Transform (DCT).
- 4) Each transformed block is quantized with respect to an  $8 \times 8$  quantization matrix, which can be chosen independently for all three color channels.
- 5) The resulting data is compressed, using Huffman or arithmetic coding.

### B. Continuous Laplacian Distribution

We collect the transformed coefficients (from Step 3 of Section II-A) for each color component by frequency. There are 64 frequencies, and it is known ([9]–[12]) that the histogram of coefficients corresponding to any one of the 63 AC frequencies resembles a *Laplacian distribution*.

Recall that the probability density function  $f(x)$  of a Laplacian with zero mean and parameter  $\lambda$  is  $f(x) = (\lambda/2)\exp(-\lambda|x|)$ , and that the corresponding cumulative distribution function  $F(x)$  is given explicitly by

$$F(x) = \int_{-\infty}^x f(t)dt = \begin{cases} \frac{1}{2} \exp(\lambda x), & \text{if } x \leq 0; \\ 1 - \frac{1}{2} \exp(-\lambda x), & \text{if } x \geq 0. \end{cases}$$

Let  $X$  be a continuous random variable with probability function  $f$ . Then the expected value of  $X$  is  $E(X) = 0$ , and the variance of  $X$  is known to be  $Var(X) = 2/\lambda^2$  [11], [12].

Comparing the distribution of the DCT coefficients to a Laplacian distribution requires care — the former is discrete, while the latter is continuous.

### C. Discrete, Dequantized Laplacian Distribution

We now focus on the distribution of the dequantized DCT coefficients. For a fixed AC frequency and a fixed color component, let  $q \in \mathbb{N}^* = \{1, 2, 3, \dots\}$  be the corresponding *quantizer* (i.e., the quantization matrix entry associated with the fixed frequency) and

$$y_1 = q \cdot \eta_1, \quad y_2 = q \cdot \eta_2, \quad \dots, \quad y_J = q \cdot \eta_J$$

be the *dequantized* DCT coefficients, i.e.: the original image is partitioned into  $J$   $8 \times 8$  blocks, and the *quantized* DCT coefficients are  $\eta_1, \eta_2, \dots, \eta_J \in \mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ . We view  $y_1, y_2, \dots, y_J$  as a random sample of size  $J$  taken by the discrete random variable

$$Y = q \cdot \text{round} \left( \frac{X}{q} \right).$$

*Proposition 1* ([13], [14]): Let  $p_q$  be the probability density function associated with the discrete random variable  $Y$ . Then

$$p_q(nq) = \begin{cases} 1 - \exp\left(-\frac{\lambda q}{2}\right), & \text{if } n = 0; \\ \exp(-\lambda q|n|) \sinh\left(\frac{\lambda q}{2}\right), & \text{if } n \in \mathbb{Z} \setminus \{0\}. \end{cases}$$

Moreover,  $E(Y) = 0$  and  $Var(Y) = (q^2 \cosh(\lambda q/2))/(2 \sinh^2(\lambda q/2))$ .  $\square$

### D. Maximum Likelihood Estimation of the Laplacian Parameter $\lambda$

Our method of choice for estimating the Laplacian parameter  $\lambda$  from the observed dequantized DCT coefficients  $y_1 = \eta_1 \cdot q, y_2 = \eta_2 \cdot q, \dots, y_J = \eta_J \cdot q$  (see Section II-C) is the following result of Price and Rabbani, whose original formulation considered the quantized coefficients  $\eta_j$ .

TABLE I  
STATISTICS OF THE ML ESTIMATES OF  $\lambda$

minimum	0.0129
1 <sup>st</sup> quartile	0.0710
median	0.1240
3 <sup>rd</sup> quartile	0.2115
maximum	1.3794
mean	0.1788
standard deviation	0.1910

*Proposition 2 (Method of Maximum Likelihood) [13], [14]:* Let  $\alpha$  be the number of indices  $j \in \{1, \dots, J\}$  with  $y_j = 0$ ,  $\beta = J - \alpha$ , and  $\rho = 2 \sum_j |\eta_j|$ . Then

$$\lambda_{\text{ML}} = \frac{-2}{q} \ln \left( \frac{\sqrt{\alpha^2 + 4(\rho - \beta)(\rho + J)} - \alpha}{2(J + \rho)} \right)$$

is the *Maximum Likelihood estimator* for  $\lambda$ . (If  $\beta = 0$ , then  $\lambda_{\text{ML}}$  becomes  $+\infty$ .)  $\square$

We considered eight standard images (bird, boat, camera, frog, goldhill, Lena, mandrill, and peppers, taken from [15]), and compressed each to quality Q100, Q90, Q80, Q70, Q60, and Q50, yielding a total of 48 test images. Using Proposition 2, we estimated the Laplacian parameter  $\lambda$  by the Method of Maximum Likelihood<sup>1</sup> for each DCT AC coefficient. The statistics, summarized in Table I, forms the basis for the numerical experiments in Section IV.

### III. DEQUANTIZATION: DIRECT VERSUS DETOUR

We now consider the problem of dequantizing twice, which occurs implicitly in the process of recompressing JPEG images by requantization.

Suppose  $q_0$  and  $q_1$  belong to  $\mathbb{N}^*$  with  $q_0 < q_1$ . Let  $x$  be in  $\mathbb{N}^*$ . Compare the *direct dequantization* of  $x$  with respect to  $q_1$ , i.e.,  $x \mapsto q_1 \cdot \text{round}(x/q_1)$ , to the *two-step dequantization* of  $x$  with respect to  $q_1$ , preceding a “detour” via  $q_0$ :  $x \mapsto q_1 \cdot \text{round}(q_0 \cdot \text{round}(x/q_0)/q_1)$ . We think here of  $q_0$  and  $q_1$  as entries of the quantization tables of the old and new requantized images, respectively, with both quantizing the same AC entry. Numerical experiments suggest that the difference between the two dequantizations cannot be described more easily except for special cases. Some basic properties are summarized in the next result.

*Proposition 3 (Direct versus Detour) ([14]):* Let  $x_1 = q_1 \cdot \text{round}(x/q_1)$ ,  $y_0 = q_0 \cdot \text{round}(x/q_0)$ , and  $y_1 = q_1 \cdot \text{round}(y_0/q_1)$ . Let  $l = \text{lcm}(q_0, q_1)$ , the least common multiple of  $q_0$  and  $q_1$ . Then:

- i)  $x_1 - y_1 \in \{-q_1, 0, +q_1\}$ .
- ii) If  $\tilde{x} = x + l$ , then  $\tilde{x}_1 = x_1 + l$  and  $\tilde{y}_1 = y_1 + l$ .
- iii) If  $m \in \mathbb{N}^*$  is odd and  $q_1 = mq_0$ , then  $x_1 = y_1$ .

<sup>1</sup>The Kolmogorov-Smirnov test shows that the Maximum Likelihood estimator is superior to estimators that are based on the Method of Moments; see [14] for details.

- iv) even multiples over estimate: let  $m \in \mathbb{N}^*$  be even,  $q_1 = mq_0$ , and  $1 \leq x \leq l$ . Then either  $y_1 = x_1 + q_1$  if

$$\frac{q_1}{2} - \left\lfloor \frac{q_0}{2} \right\rfloor \leq x \leq \frac{q_1}{2} - 1$$

or  $y_1 = x_1$ , otherwise.

- v) Let  $m \in \mathbb{N}^*$  be even,  $q_1 = 1 + mq_0$ , and  $1 \leq x \leq l$ . Then  $y_1 = x_1$ , if  $x < (q_1 + 1)/2$ ; and  $y_1 = x_1 - q_1$ , if

$$\frac{q_1 + 1}{2} \leq x \leq \frac{q_1 - 3}{2} + \left\lceil \frac{q_0}{2} \right\rceil.$$

The error for larger  $x$  is more complicated.  $\square$

The *overestimation* in Proposition 3.(iv) comes from the fact that the “round” function rounds positive half-integers *up*. For general  $q_0 < q_1$ , the problem of concisely and usefully determining the difference  $x_1 - y_1$ —without actually computing it—is quite difficult. Therefore, we will discuss numerical approaches to this problem in Section IV.

### IV. DEQUANTIZATION: ERROR AND ENTROPY FOR THE LAPLACIAN

Since the distribution of each DCT AC coefficient is well approximated by a Laplacian (Section II-C), recompressing a JPEG image by requantization leads to questions about the behavior of the error resulting from dequantizing a Laplacian twice. In this section, the problem of determining the error is made mathematically precise and tackled numerically. We also consider the entropy of the twicely dequantized Laplacian.

#### A. Mathematical Description

Fix a Laplacian parameter  $\lambda > 0$ , and  $q_0, q_1 \in \mathbb{N}^*$ . The error made by dequantizing  $n \in \mathbb{Z}$  in two steps is given by

$$e_n = q_1 \cdot \text{round} \left( q_0 \cdot \frac{\text{round} \left( \frac{n}{q_0} \right)}{q_1} \right) - n.$$

Since the dequantization convention in JPEG [5] is defined in terms of nonnegative values, it suffices to consider the case when  $n \geq 0$ . Employing Proposition 1 (with  $q = 1$ ) and its notation, we see that the (*weighted absolute*) error made by dequantizing the given Laplacian in two steps is

$$e = 2 \sum_{n \in \mathbb{N}^*} p_1(n) |e_n|.$$

Our definition of  $e_n$  is simpler than that of Chan [6], which compares the outcome of dequantizing twice to the result of dequantizing directly; we believe our approach more closely resembles the actual requantization environment, as one typically wants to be close to the original. However, we employ her key idea of breaking up the error into two components: we write our error  $e$  as  $e = e_+ + e_-$ , where  $e_+$  and  $e_-$  are both nonnegative and measure *increase* and *decrease* in amplitude, respectively. The components are given by

$$e_+ = 2 \sum_{n \in \mathbb{N}^*: e_n > 0} p_1(n) e_n \quad \text{and} \\ e_- = 2 \sum_{n \in \mathbb{N}^*: e_n < 0} p_1(n) |e_n|.$$

Note that  $e, e_+, e_-$  all depend implicitly on  $\lambda, q_0, q_1$ . With these definitions, we can now interpret the two-stage dequantization

of the Laplacian within the context of JPEG images: *If the component  $e_+$  (resp.  $e_-$ ) dominates, then the amplitude increases (resp. decreases) and we thus expect the image dequantized in two steps to be grainier/sharper (resp. blockier/smoothier) than the original.*

Keeping  $\lambda$  and  $q_0$  fixed, one would expect that these errors are increasing functions of  $q_1$ . Surprisingly, this is not the case; hence we ideally aim to find *local minimizers*, and in this way the problem of requantizing thus amounts to finding  $q_1 \geq q_0$  such that  $q_1$  is a local minimum of these error measures. As this appears to be intractable analytically, we will reason numerically.

It will be illuminating to track the entropy of the doubly dequantized Laplacian.<sup>2</sup> Let  $X$  be as in Section I, and define discrete random variables

$$Y = q_0 \cdot \text{round} \left( \frac{X}{q_0} \right) \quad \text{and} \quad \tilde{Y} = q_1 \cdot \text{round} \left( \frac{Y}{q_1} \right).$$

Proposition 1 (with  $q = q_0$ ) provides an explicit formula for the probability density function corresponding to  $Y$ . While the probability density function corresponding to  $\tilde{Y}$ , which we denote by  $\tilde{p}$ , admits no closed form, it can easily be computed numerically. The *entropy* of the corresponding doubly dequantized Laplacian is now defined by

$$h = - \sum_{n \in \mathbb{Z}} \tilde{p}(n) \ln(\tilde{p}(n))$$

where  $0 \cdot \ln(0) = 0$  by convention. Note that  $h$  depends implicitly on  $\lambda, q_0, q_1$ . Analogously to the preceding discussion of the error, we are interested in the behavior of  $h$ —where  $\lambda$  and  $q_0$  are fixed but  $q_1 \geq q_0$  varies—as a measure of compressibility.

### B. Numerical Experiments: Approach and Results

To study numerically  $e, e_+, e_-$ , and  $h$  as functions of  $q_1$ , we require a sensible range for the Laplacian parameter  $\lambda$ , and a suitable point of truncation. In view of Table I, we choose the range of  $\lambda$  to be  $[0.01, 2]$ , while the point of truncation is selected to conform with the following fact that the JPEG image format implicitly constrains the range of DCT coefficient values:

*Proposition 4 [14]:* Suppose  $X$  is an  $8 \times 8$  integer matrix with values in  $\{0, 1, 2, \dots, 255\}$ , and let  $Y$  be the two-dimensional DCT of  $X$ . Then the magnitude of every AC DCT coefficient is never larger than 1020.  $\square$

Therefore, we considered various values of  $\lambda \in [0.01, 2]$ ; for the corresponding Laplacian distribution with frequency  $p_1(n)$  (see Proposition 1 with  $q = 1$ ), we truncated the distribution at 1020 and renormalized.

We computed the errors of Section IV-A numerically, for all  $q_0, q_1$  in  $\mathbb{N}^*$  with  $1 \leq q_0 \leq q_1 \leq 255$ , corresponding to the JPEG baseline standard. We then collected the errors in an upper triangular  $255 \times 255$  matrix, where the  $(q_0, q_1)$  entry is the error corresponding to the fixed  $\lambda$ . The resulting matrices of errors  $e_+, e_-, e$ , respectively, are visualized in Fig. 4–6, respectively, for  $\lambda = 0.0710$  (first quartile), chosen because the important features are more pronounced and hence easier to spot.<sup>3</sup>

<sup>2</sup>We are grateful to an anonymous referee for making this pertinent suggestion.

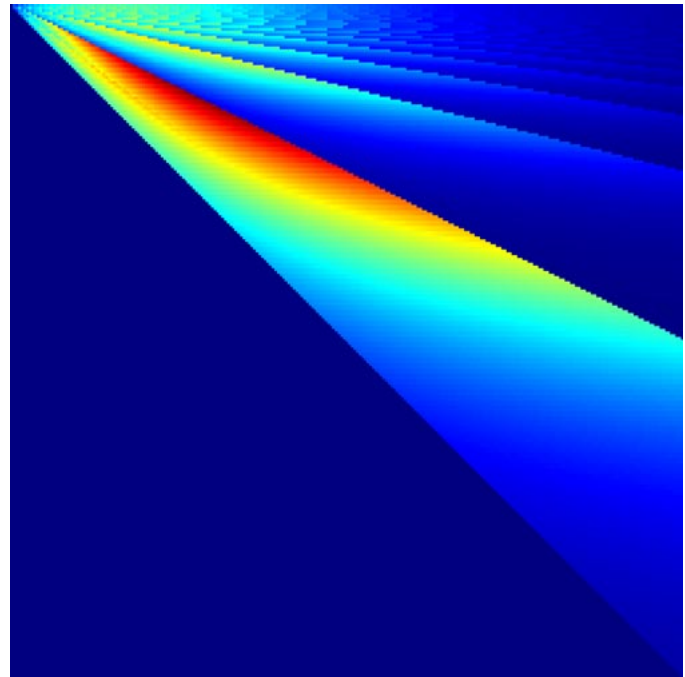


Fig. 4.  $\lambda = 0.0710$  (first quartile), error  $e_+$ , maximum entry 11.8.

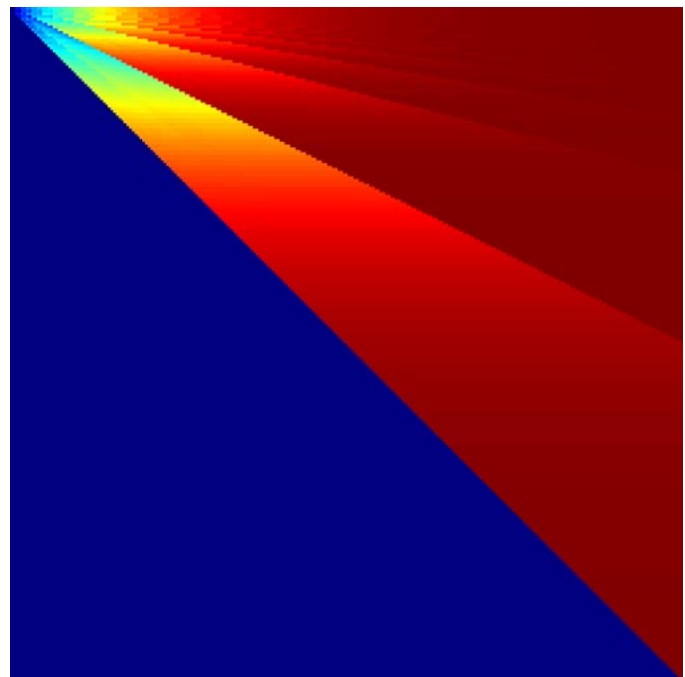


Fig. 5.  $\lambda = 0.0710$  (first quartile), error  $e_-$ , maximum entry 14.1.

When these error matrices are calculated using a different distribution, quite different behavior is observed, which underlines the crucial importance of the Laplacian distribution to this phenomenon. These images—along with qualitatively similar error images corresponding to other values of  $\lambda$ , which we omit for brevity—give rise to the following observations:

<sup>3</sup>The matrices were created with GNU Octave [16]. We used the color map `jet`, with range: darkblue (minimum)  $\rightarrow$  blue  $\rightarrow$  cyan  $\rightarrow$  green  $\rightarrow$  yellow  $\rightarrow$  red  $\rightarrow$  dark red (maximum). To bring out the features for  $e_+$  and  $e_-$  more clearly, we displayed the element-wise square root of these error matrices.

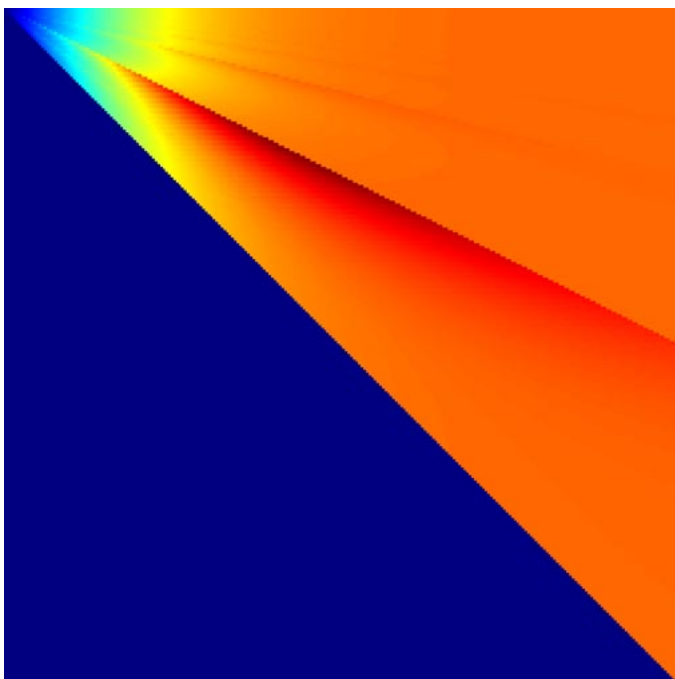


Fig. 6.  $\lambda = 0.0710$  (first quartile), error  $e = e_+ + e_-$ , maximum entry 18.2.

- the  $e$  error images exhibit ridges when  $q_1$  is approximately an even multiple of  $q_0$  (see Fig. 6);
- the  $e_+$  (resp.  $e_-$ ) error images exhibit the same ridges (see Fig. 4 resp. Fig. 5);
- this behavior is gradual as  $\lambda$  increases, with the ridges disappearing more quickly.

Figs. 4– 6 give us a good view of the overall behavior. It is instructive to inspect a fixed *row* of the error matrices, i.e., the error curves for fixed  $\lambda$ ,  $q_0$  and varying  $q_1$ , and to observe the striking split of  $e$  into the components  $e_+$  and  $e_-$ . See Fig. 7, where  $\lambda = 0.1788$  (mean) and  $q_0 = 5$ . We point out that the characteristics of fixed *rows* and also fixed *columns* of the error matrices can be used to explain requantization phenomena reported by Chan [6] for quotients  $q_1/q_0$  close to 2, 3, or 4.

Handling the entropy  $h$  analogously, we conclude that

- the entropy  $h$  decreases to 0, with marked drops when  $q_1$  is an even multiple of  $q_0$ .

See Fig. 8, where  $\lambda = 0.1788$  (mean) and  $q_0 = 5$  (which are the same parameters as those used in the generation of Fig. 7).

## V. REQUANTIZATION ERROR FOR JPEG IMAGES

### A. Key Observations

The numerical results of Section IV-B lead us to postulate the following.

*Observation 5:* Suppose  $\lambda$  and  $q_0$  are fixed, and denote the set of positive even integers by  $M$ . Then  $e$ ,  $e_+$ ,  $e_-$ ,  $h$ , viewed as functions of  $q_1$ , satisfy the following.

- The total error  $e$  is monotone increasing, with the exception of “dippers” around points  $q_1 = mq_0 + 1$ , where  $m \in M$ .
- The (amplitude-intensifying) error  $e_+$  is overall decreasing to 0. More importantly, it is increasing on

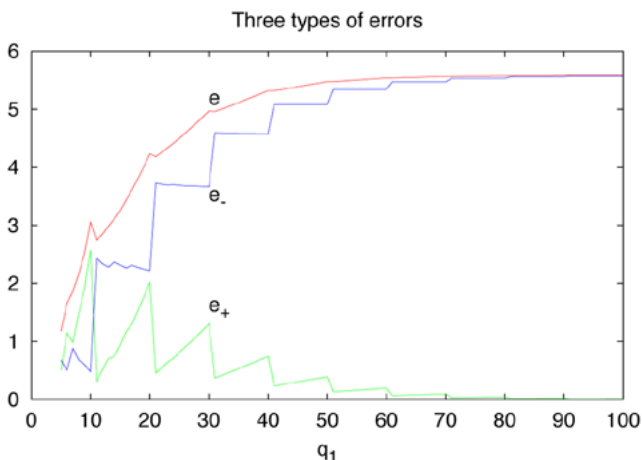


Fig. 7.  $\lambda = 0.1788$  (mean),  $q_0 = 5$ . Red ( $e$ ), Green ( $e_+$ ), Blue ( $e_-$ ).

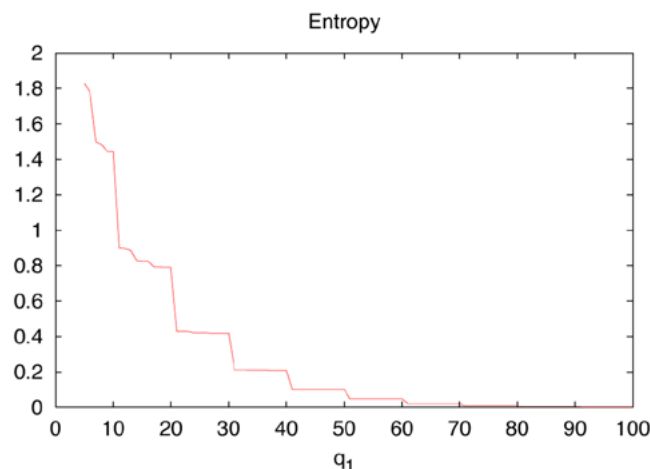


Fig. 8.  $\lambda = 0.1788$  (mean),  $q_0 = 5$ , entropy  $h$ .

intervals  $[mq_0 + 1, (m + 2)q_0]$ , and there is a sharp drop between  $mq_0$  and  $mq_0 + 1$ , where  $m \in M$ .

- The (amplitude-reducing) error  $e_-$  is overall increasing and approaching  $e$ . More significantly, it is nearly constant on intervals  $[mq_0 + 1, (m + 2)q_0]$ , with a sharp rise between  $mq_0$  and  $mq_0 + 1$ , where  $m \in M$ .
- The entropy  $h$  (a measure of compressability) is decreasing to 0, with a particularly sharp drops between  $mq_0$  and  $mq_0 + 1$ , where  $m \in M$ .

Note that the sharp changes of  $e_+$  (resp.  $e_-$ ) in (ii) (resp. (iii)) of Observation 5 parallel the errors of two-step versus direct requantization in item (iv) (resp. item (v)) of Proposition 3. Observation 5 has the following consequences for recompressing JPEG images by requantization.

*Observation 6 (consequences for Recompression):* The quotient  $q_1/q_0$  is crucial:

- there is a sharp divide in the type of error and in the rate of compression for values of  $q_1$  in the vicinity of an even multiple of  $q_0$ , say  $q_1 \approx mq_0$ , where  $m \in \mathbb{N}^*$  is even;
- recompressing with  $q_1 = mq_0$  or less is amplitude intensifying and thus sharper/grainier images are obtained;
- recompressing with  $q_1 = mq_0 + 1$  or more is amplitude reducing and so smoother/blockier images are received;

- *recompression with  $q_1 = mq_0 + 1$  or more yields significantly better compression rates than recompression with  $q_1 = mq_0$  or less.*

### B. An Example: The Introductory Images Revisited

Recall that the image given in Fig. 1 is of quality Q75. By Observation 6, the crucial quantity is the element-wise quotient of the new quantization matrix divided by Q75. It turns out that all entries of these element-wise quotient matrices Q50./Q75 and of Q48./Q75 are very close to 2: specifically, 55.6% and 12.7% of Q50./Q75 and of Q48./Q75, respectively, are equal to 2, with 44.4% of Q50./Q75 slightly less than 2 and 84.1% of Q48./Q75 slightly greater than 2.

*In view of Observation 6, and compared to the given image shown in Fig. 1, we thus predict the image shown in Fig. 2, corresponding to Q50, to be much sharper/grainier, and the image shown in Fig. 3, corresponding to Q48, to be much smoother/blockier. This is precisely the case; in fact, we perceive the Q48 version as smooth, and the Q50 version as grainy. Most importantly, the Q48 version appears to be perceptually closer to the original than the Q50 version. Moreover, as predicted and already pointed out in the Introduction, the Q48 version is substantially smaller in file size than the Q50 version.*

## VI. CONSEQUENCES FOR RECOMPRESSION

We now present a heuristic algorithm for recompressing JPEG images that avoids the aforementioned “grainy” artifacts and seeks “smooth” artifacts instead. Proposition 3 and Section IV-B show that the situation where the new quantizer  $q_1$  is (close to) an integer multiple of the old quantizer  $q_0$ , say  $q_1 = mq_0$  for some  $m \in \mathbb{N}^*$ , is special: in fact, if  $m$  is odd, then the “detour” via  $q_0$  does not affect the resulting image (Proposition 3.(iii)). However, if  $m$  is even, then recompression is very sensitive, with strong amplitude increase (resp. decrease) for  $q_1 = mq_0$  (resp.  $q_1 = mq_0 + 1$ ).

Now suppose we need to recompress a given (already compressed) JPEG image. Assume  $Q_0$  (resp.  $Q_{\text{bound}}$ ) is the original (resp. target) quantization matrix, respectively. Using these two matrices, we wish to find a new quantization matrix  $Q_1$ , in the vicinity of  $Q_{\text{bound}}$ , in order to requantize the given image. For a fixed frequency, denote the corresponding quantizers by  $q_0$ ,  $q_{\text{bound}}$ , and  $q_1$ , respectively.

### A. Heuristic Algorithm (Preliminary Version)

The new  $q_1$  is constructed as follows: Let  $k = \lfloor q_{\text{bound}}/q_0 \rfloor$ . Then define

$$q_1 = \begin{cases} q_0, & \text{if } k = 0; \\ kq_0, & \text{if } k \text{ is odd;} \\ kq_0 + 1, & \text{if } k > 0 \text{ and } k \text{ is even.} \end{cases}$$

We do this for each of the 64 frequencies to obtain a new quantization matrix  $Q_1$ , which is then used to recompress the given image. The new quantization matrix picks up the little “dips” of  $e$  for every frequency, which are also local minima for  $e_+$ .

By the very nature of its construction, this algorithm is good at avoiding “grainy” recompression artifacts, but it is unclear how it could achieve predetermined compression rates exactly.

An interesting topic for further research is the design of an algorithm that would not only possess the perceptual qualities of the present algorithm but that would also allow to specify the desired compression rate in advance.

### B. Heuristic Algorithm (Implemented and Experimentally Validated Version)

In Section VI-A, we had to consider separate cases ( $k$  odd and  $k$  even) to avoid the increase in amplitude. However, it is possible to unify these cases by modifying the rounding convention used in the requantization, resulting in the following heuristic algorithm:

First compute  $k$  as in Section VI-A:  $k = \lfloor q_{\text{bound}}/q_0 \rfloor$ . Secondly, define  $q_1$  by

$$q_1 = \begin{cases} q_0, & \text{if } k = 0; \\ kq_0, & \text{if } k > 0. \end{cases}$$

Thirdly, by treating all frequencies, assemble a new quantization matrix  $Q_1$ . Finally, requantize the given image with respect to  $Q_1$  but observe a *modified rounding convention*, which rounds genuine positive half-integers ( $1/2, 3/2, \dots$ ) down (toward zero).

We performed subjective tests to validate our predicted results: we asked nine subjects to compare the implemented heuristic algorithm of Section VI-B to a blind requantization using the *Mean Opinion Score* rating from [17]. For all subjects, the heuristic algorithm performed at least as well as a blind requantization over most images, but noticeable improvements in the subject’s ratings were observed for images in which the aforementioned nonmonotonicity in the IJG quality rating scale had been observed, thus confirming our predicted results. A detailed discussion of our experiments can be found in [14].

## VII. CONCLUSION

In this paper, we consider the problem of requantizing JPEG images. We formulate and study this problem by utilizing the Laplacian distribution of the DCT AC coefficients and by splitting the error made during requantization into amplitude-increasing and amplitude-decreasing components. Our analysis explains the surprising perceptual nonmonotonicity of the IJG quality scale and also the apparent content independence of some requantization phenomena reported by Chan [6]. Moreover, it leads to an experimentally validated algorithm for finding new quantization matrices for recompression.

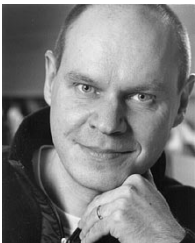
### ACKNOWLEDGMENT

The authors wish to thank the referees and the associated editor for their helpful comments and suggestions.

### REFERENCES

- [1] “Various Documents and Links,” Joint Photographic Experts Group, <http://www.jpeg.org>.
- [2] W. B. Pennebaker and J. L. Mitchell, “JPEG Still Image Data Compression Standard,” Van Nostrand Reinhold, New York, 1993.
- [3] J. Miano, *Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP*. New York: ACM Press, 1999.
- [4] D. Hankerson, G. A. Harris, and P. D. Johnson Jr., *Introduction to Information Theory and Data Compression*. Boca Raton, FL: CRC, 1997.
- [5] “Various Documents,” Independent JPEG Group, <http://www.ijg.org>.

- [6] S. Chan, "Recompression of Still Images," University of Kent, Canterbury, U.K., Tech. Rep. 2-92\*, 1992.
- [7] —, "The Use of the JPEG Image Compression Standard and the Problem of Recompression," Ph.D. dissertation, University of Kent, Canterbury, U.K., 1992.
- [8] A. Bovik, *Handbook of Image and Video Processing*, A. Bovik, Ed. New York: Academic, 2000.
- [9] R. C. Reininger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM-31, pp. 835–839, June 1983.
- [10] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Processing*, vol. 9, pp. 1661–1666, Oct. 2000.
- [11] S. R. Smoot and L. A. Rowe, "Laplacian model for AC DCT terms in image and video coding," *Proc. 9th IEEE Image and Multidimension Digital Signal Processing Workshop*, 1996.
- [12] —, "Study of DCT coefficient distribution," in *Human Vision and Electronic Imaging*, B. E. Rogowitz and T. N. Pappas, Eds. Bellingham, WA: SPIE, 1996, vol. 2657.
- [13] J. R. Price and M. Rabbani, "Biased reconstruction for JPEG decoding," *IEEE Signal Processing Lett.*, vol. 6, pp. 297–299, Dec. 1999.
- [14] M. S. Macklem, "Multidimensional Modeling of Image Fidelity Measures," M.Sc. thesis, Dept. Mathematics, Simon Fraser University, Burnaby, BC, Canada, 2002.
- [15] *Standard Test Images*, <http://links.uwaterloo.ca/bragzone.base.html>.
- [16] J. W. Eaton *et al.*, "GNU Octave," <http://www.octave.org>.
- [17] J. Allnatt, *Transmitted-Picture Assessment*. New York: Wiley, 1983.



**Heinz H. Bauschke** (M'00) received the "Diplom-Mathematiker (mit Auszeichnung)" degree in 1990 from Goethe University, Frankfurt, Germany, and the Ph.D. degree in mathematics in 1996 from Simon Fraser University, Burnaby, BC, Canada, where he graduated with the Governor General's Gold Medal for achieving the highest academic standing at the graduate level.

From 1996 to 1998, he was an NSERC Postdoctoral Fellow at University of Waterloo, University of Santa Barbara, and Pennsylvania State University. He then spent 3 years in Kelowna, BC, teaching at Okanagan University College and researching for Packeteer Canada, where most of this work was completed. He is now Assistant Professor at the University of Guelph, Ontario, in the Department of Mathematics and Statistics. His current research interests include optimization, image processing, and compression.



**Christopher H. Hamilton** (S'02) received the B.Sc. degree in mathematics and physics in 2001 at Okanagan University College in Kelowna, BC, Canada. He is currently pursuing the M.S. degree at Simon Fraser University, Burnaby, BC, under the supervision of Dr. Jonathan Borwein.

He is also a Research Associate at Packeteer Canada, where he works on data/image compression and image processing related problems.



**Mason S. Macklem** (S'01) received the B.A. degree in mathematics in 2000 from Okanagan University College, and the M.S. degree at Simon Fraser University (SFU), Burnaby, BC, Canada. He is currently pursuing the Ph.D. degree at SFU, where he is involved with the Centre for Experimental and Constructive Mathematics (CECM).

From April 2000 to August 2002, he was a Research Assistant at Packeteer Canada, focusing on perceptual image quality assessment. He is also a Research Assistant with the CoLab collaborative environment at SFU, working on developing interfaces for derivative-free optimization algorithms. His research interests include nonsmooth optimization, image compression, vector quantization, pattern recognition, and human vision modeling.

Mr. Macklem is a student member of the Society of Photonics and Interface Engineers.

**Justin S. McMichael** received the B.Sc. degree in physics in 2000 from Okanagan University College. He plans to complete an advanced degree in physics.

He has worked as a Research Associate and Engineer for Workfire, which was acquired subsequently by Packeteer. He is now working as director of advanced technology for Itiva, Inc. His responsibilities include building tools to perform the visualization, data acquisition, and basic analysis of images for this research topic.

**Nicholas R. Swart** received the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, in 1995.

After graduating, he spent two years as a Researcher at the National Optics Institute, Sainte-Foy, QC, Canada, working on design and numerical modeling of micromachined optical sensors and actuators. He subsequently spent three years at Analog Devices, Cambridge, MA, where he worked as a CAD Manager and microsystem design engineer. In addition, he managed a large software research project in the area of sensor modeling, design and optimization, which was funded with a multimillion dollar grant from DARPA. More recently, he was Vice-President of Advanced Technology of Workfire Technologies Inc., a company which developed networking software, and which was acquired in September of 2000 by Packeteer, Inc., leading to his position as Director of Advanced Technology. He is now with Analog Devices as a Senior CAD Engineer.