

ON TRANS-CODING OPTIMIZATION USING RE-QUANTIZATION

O. Gendler and M. Porat

Department of Electrical Engineering Technion - Israel Institute of Technology, Haifa 32000, Israel <u>orag@tx.technion.ac.il</u>, <u>mp@ee.technion.ac.il</u>

ABSTRACT

In this work, re-quantization for trans-coding of MPEG intra-frames and JPEG images is considered and analyzed. Our analysis shows that both the rate and the distortion of re-quantized images depend mainly on the ratio between the new and the old quantization steps. The new quantization step is selected using a simplified fast algorithm that ensures low distortion. Our analysis is based on the structure of the quantizer and the Laplace-like distribution of the DCT coefficients in sub-band coding. The proposed approach could be instrumental in achieving a required bit-rate at low distortion while allowing real-time implementation due to low computational complexity.

Key Words: Image Coding, Re-quantization, Trans-rating, Trans-coding, Rate-Distortion, Computational complexity.

1. INTRODUCTION AND PROBLEM DEFINITION

Image and video transmission usually requires trans-coding since in many cases the bit-rate of the transmitted data has to meet the various requirements of the transmission channel and the limitations of the receiving end. This process of bit-rate reduction has to be carried out with minimal distortion, and when done in real-time, it is crucial that the trans-coding is performed at low computational complexity. A straightforward method is to cascade a decoder and an encoder. Despite the simplicity of this approach, however, such a system is computationally heavy and not necessarily optimal. More sophisticated trans-coding schemes can be performed in many ways, both in the bit domain and in the frequency domain [1].

One of the methods for video trans-rating is re-quantization. This is done by re-quantizing DCT coefficients in the transform domain. Figure 1 depicts general system description. Figure 1(a) shows a basic coding scheme where DCT is performed on each image block. The DCT coefficients are then quantized and coded using variable length coding (VLE). Figure 1(b) shows a basic re-quantization scheme. The coded signal is decoded (VLD), re-quantized in the DCT domain and then coded again (VLE). Two stages of quantization are involved in the process of re-quantization. The first stage of quantization is performed at the encoder and cannot be controlled, altered or avoided. It is shown as the Q_1 block in Figure 1 and the quantization step size used is q_1 . The second stage of quantization is performed for trans-coding or trans-rating. It is shown as the Q_2 block in Figure 1 and the quantizer, a third reference quantizer is also used throughout this work denoted by $Q_{2,ref}$. The reference coarse quantizer is used directly on the original signal. The performance of the second stage quantizer is evaluated by comparing it to the reference quantizer.

The goal of this work is to analyze the process from a rate-distortion point of view, so that it is possible to design this second stage quantizer in an optimal way that would keep the distortion low as well as the amount of bits required to code the data. The developed method has to be fast and simple so that it could be suitable for real-time applications.



Figure 1: (a) Basic coding scheme; (b) Basic re-quantization scheme

1.1. Quantizer Definitions

Typically, in both MPEG Intra frames and JPEG images, a uniform threshold quantizer (UTQ) is used, with no dead zone. Such a quantizer (also called a midtread quantizer) as illustrated in Figure 2 is used for both quantization stages, as in [3] and [5].

Throughout this work two types of quantization are considered. The first is uniform quantization, i.e., the entire image is quantized with the same quantization step. The second is quantization using the JPEG quantization matrix when only the quality factor Q is altered. This factor multiplies the constant quantization matrix. The quantizer is well defined using Equations (1) and (2), similar to definitions used in [5].



Figure 2: UTQ schematic description

$$y_{1} = Q_{1}(x) = \lfloor x/q_{1} + 0.5 \rfloor q_{1}$$

$$y_{2} = Q_{2}(Q_{1}(x)) = \lfloor \lfloor x/q_{1} + 0.5 \rfloor q_{1}/q_{2} + 0.5 \rfloor q_{2}$$

$$y_{2,ref} = Q_{2,ref}(x) = \lfloor x/q_{2} + 0.5 \rfloor q_{2}$$
(1)
$$(Q(x) = x \ge 0)$$

$$y = \begin{cases} Q(x) &, x \ge 0\\ -Q(|x|), x < 0 \end{cases}$$
(2)

The decision and reconstruction levels of the quantizer are defined as in Figure 2. The decision level l of the stage i (i=1,2) quantizer is denoted by $d_{i,l}$. In general, $d_{i,0} = 0$ and $d_{i,l} = (l - 0.5)q_i$. The reconstruction level l of stage i quantizer is denoted by $r_{i,l}$. When using a midtread quantizer, $r_{i,0} = 0$ and $r_{i,l} = l \cdot q_i$. Some formulas were developed for positive values of the DCT coefficients. For negative values (2) is used.

This work is organized as follows. In Section 2, a method for selecting a new quantization step to minimize the MSE is shown. Section 3 introduces theoretical rate-distortion analysis. In Section 4 simulation results are presented, and in Section 5 the work is summarized and conclusions are provided.

2. MINIMZING RE-QUANTIZATION DISTORTION

The criterion used here for distortion is MSE (Mean Squared Error) as well as visual quality as discussed in Section 4. Consider all possible input values of one bin (bin l) of the first stage quantizer Q_l as

$$x \in [d_{1,l}, d_{1,l+1}) \xrightarrow{Q_1} r_{1,l} \equiv const \xrightarrow{Q_2} r_{2,m} \equiv const,$$
(3)

as suggested in [3], in order to keep the re-quantization distortion to a minimum, we should attempt to achieve

$$y_2 = Q_2(Q_1(x)) = y_{2,ref} = Q_{2,ref}(x).$$
(4)

This way, there is no additional distortion caused by using two- stage quantization instead of using one stage coarser reference quantizer directly. Specifically we can write

$$x \in [d_{1,l}, d_{1,l+1}) \xrightarrow{Q_{2,ref}} r_{2,m} \equiv const$$
(5)

$$\forall l, \exists m: [d_{1,l}, d_{1,l+1}) \in [d_{2,m}, d_{2,m+1}) .$$
(6)

This means, that in order to avoid a re-quantization error, each quantization bin of the first stage quantizer Q_1 has to be contained in some quantization bin of the coarser second stage quantizer Q_2 . Otherwise, if there is some second stage decision level $d_{2,m}$ between (and not equal to) a pair of first stage decision levels ($d_{2,m} \in [d_{1,l}, d_{1,l+1})$) and $d_{2,m} \neq d_{1,l}, d_{1,l+1}$), condition (4) is contradicted. This is shown in the following equations, assuming for instance that $d_{2,m} < r_{1,l}$.

$$x \in [d_{1,l}, d_{1,l+1}) \xrightarrow{Q_1} r_{1,l} \xrightarrow{Q_2} r_{2,m}$$

$$x \in [d_{2,m}, d_{2,m+1}) \xrightarrow{Q_{2,ref}} r_{2,m}$$

$$x \in [d_{2,m-1}, d_{2,m}) \xrightarrow{Q_{2,ref}} r_{2,m-1}$$

$$\exists x \in [d_{1,l}, d_{2,m}) : Q_2(Q_1(x)) \neq Q_{2,ref}(x)$$

$$(8)$$

This case is illustrated in Figure 3.



Figure 3 : Quantization bin illustration

In order to avoid additional distortion due to re-quantization and maintain the condition in (4), from (6), (7) and (8) it follows that

$$\left\{d_{2,m}\right\} \subseteq \left\{d_{1,l}\right\}.$$
(9)

This means that all decision levels of the second stage quantizer must coincide with decisions levels of the first stage quantizer. Thus, each bin of the coarse second stage quantizer must contain an integer multiplication of the finer first stage quantizer bins: $q_2 = k \cdot q_1$, where k is a natural number.

Consider the first bin of the coarse reference quantizer: $d_{2,ref,0}=0$; $d_{2,ref,1}=0.5q_2$. When taking into account the quantization bins of the first stage quantizer that fits into the large reference bin we should use

$$d_{1,0} = 0;$$
 $d_{1,n} = (n - 0.5)q_1$, (10)

where $n(n \ge l)$, is the index of the decision level of the finer quantizer that coincides with the $d_{2,ref,l}$ decision level of the coarse reference quantizer. We may also state: $d_{2,ref,l} = d_{1,n}$. Accordingly, we get

$$0.5 \cdot q_2 = (n - 0.5) \cdot q_1$$

$$q_2 / q_1 = (n - 0.5) / 0.5 = k$$
(11)

$$k = 2n - 1 = odd \tag{12}$$

This shows that in order to maintain the condition in (4) for all the x axis, the ratio between the finer and the coarser quantizer steps, $k = q_2/q_1$ must be an odd number.

Further analysis of the distortion due to re-quantization appears in the following section.

3. RATE-DISTORTION ANALYSIS FOR LAPLACIAN PDF

Analysis of the distortion and bit rate of the re-quantized image requires modeling the distribution of the DCT coefficients throughout the entire re-quantization process. First, the original data has to be modeled, next the quantized data and finally the re-quantized data. The bit rate of the re-quantized image could then be estimated for various re-quantization step sizes by calculating the entropy of the re-quantized data. The optimal re-quantization steps, which minimize the entropy (and thus the bit-rate) and the distortion of the re-quantized data can then be selected.

DCT coefficients of coded images can be fairly modeled using Laplace distribution [3],[4],[9] such that $p(x) = 0.5\lambda e^{-\lambda|x|}$ (13)

The probability weight of each quantization bin of the first stage quantizer can be then calculated, using the quantization step size (which is known to the device performing the re-quantization). For quantization step q, for positive values, the probability weights of the first stage quantizer are expressed as

$$0.5w_{0} = 0.5\lambda \int_{0}^{0.5q} e^{-\lambda x} dx = 0.5e^{-\lambda x} \Big|_{0}^{0.5q} = 0.5 \left(1 - e^{\frac{-\lambda q}{2}}\right)$$

$$w_{1} = 0.5\lambda \int_{0.5q}^{1.5q} e^{-\lambda x} dx = -0.5e^{-\lambda x} \Big|_{0.5q}^{1.5q} = 0.5 \left(-e^{\frac{-3\lambda q}{2}} + e^{\frac{-\lambda q}{2}}\right) = 0.5e^{\frac{-\lambda q}{2}} \left(1 - e^{-\lambda q}\right)$$

$$w_{2} = 0.5\lambda \int_{1.5q}^{2.5q} e^{-\lambda x} dx = -0.5e^{-\lambda x} \Big|_{1.5q}^{2.5q} = 0.5 \left(-e^{\frac{-5\lambda q}{2}} + e^{\frac{-3\lambda q}{2}}\right) = 0.5e^{\frac{-3\lambda q}{2}} \left(1 - e^{-\lambda q}\right)$$

$$w_{l} = 0.5\lambda \int_{(l-0.5)q}^{(l+0.5)q} e^{-\lambda x} dx = -0.5e^{-\lambda x} \Big|_{(l-0.5)q}^{(l+0.5)q} = 0.5 \left(-e^{-\lambda (l+0.5)q} + e^{-\lambda (l-0.5)q}\right) = 0.5e^{-\lambda (l-0.5)q} \left(1 - e^{-\lambda q}\right)$$

$$(14)$$

The probability weight of bin l is denoted by w_l . Figure 4 depicts the probability weights of the first stage quantizer. Only the positive axis is shown and thus only half of the bin around zero appears.



Figure 4 : Probability weights after quantization

Next, it is needed to calculate the entropy and MSE of the re-quantized data in order to estimate the required bit-rate and distortion. The probability distribution at the input of the second stage quantizer is no longer Laplacian and the quantized data takes discrete values. Each value is a reconstruction level of a quantization bin from the first stage quantizer with probability weight accordingly. The re-quantized data will also obtain discrete values, according to the new reconstruction levels. In fact, it is possible to regard the re-quantization process as *rearranging of the probability weights* of the discrete quantized values into new groups, representing the re-quantized values and their probability distribution.

It can be observed that the position of the decision levels of the second stage quantizer compared to the decision levels of the first stage quantizer, determines the re-grouping of the probability weights of the quantized data. This determines the probability weights (and distribution) of the resulting re-quantized data. Detailed rate-distortion analysis using this notion is given in the following sub sections. In addition, the rounding policy of the second stage quantizer (i.e., whether 0.5 is rounded towards 1 or 0) greatly affects the re-grouping of the probability weights during the re-quantization process. Rounding 0.5 to 0 was suggested in [10] and is regarded as "rounding toward zero" while rounding 0.5 to 1 is regarded as "regular rounding".

Figure 5 demonstrates the re-quantization process and the affect of the ambiguity when rounding 0.5 towards 1 or 0. The first (left most) axis represents possible values at the input of the first stage quantizer. Decision levels are marked with longer lines perpendicular to the axis. Representation levels are marked with the shorter lines with "x" on them. The original quantization step size is 3 for all the second stage quantizers depicted, as shown in the left lower corner of the figure.

The following axes demonstrate the re-quantization process for different quantization step sizes used for the second stage quantizer. For instance, the second axis from the left represents re-quantization using a quantizer with step size equal to 4. The decision and representation levels of the second stage quantizers are marked in a similar way. Also, the transition of each representation level of the first stage quantizer to the appropriate representation level of the second stage quantizer is shown. For that purpose the following notation is used in the figure.



Let us observe closely the case when the quantization step of the second stage quantizer is equal to 6 (twice the original quantization step). It can be noticed that there is some ambiguity regarding the quantization of the value 3 (a representation value of the first stage quantizer). This value falls directly on the decision level of the second stage quantizer. This means that it is possible to quantize this value to 0 (rounding towards zero) or 6 (regular rounding). In a similar way, -3 could be quantized towards 0 or -6.

This situation leads to the conclusion that the rounding policy of the second stage quantizer can greatly affect the re-quantization process outcome. While both possible approaches to the rounding policy were considered throughout this work, rounding towards zero appears to be a more promising approach. This can intuitively be explained as follows:

Let $p(x = \alpha)$ denote the probability of a DCT coefficient x being equal to α in the original image. Since the Laplace PDF is zero centered, for $\alpha_1 > \alpha_2 > 0$: $p(x = \alpha_2) > p(x = \alpha_1)$. It thus follows that rounding to a lower value will decrease the distortion. In addition, the more values of the quantized DCT coefficients are rounded towards zero the bigger the probability weight of the zero quantization bin becomes. This affects the entropy directly and decreases the required bit-rate. Detailed analysis appears in the following sub sections. Both rounding methods were implemented and compared in the simulations described in Section 4.



Figure 5 : *Re-quantization process example*

3.1 Rate Analysis of a Laplacian Source

As mentioned, entropy will be used as a measure of the resulting bit-rate of the re-quantized data. The probability distribution of the input signal (at the first stage quantizer) is Laplacian, as defined in (13) and shown in Figure 6.



Figure 6 : Laplacian probability distribution

Three methods of re-quantization (one for each rounding method and direct quantization) were considered. Below is a derivation of the entropy as a function of the factor k, which is the ratio between the first and second stage quantization step sizes as defined by

$$k = \frac{q_2}{q_1} \,. \tag{15}$$

Only integer values are considered as shown in (9) since an integer ratio is required to avoid additional distortion due to the re-quantization process and since it allows for a comprehensive analysis regardless of the first stage (initial) quantization step size. Moreover, it is expected that for odd values of k both the rate and the distortion performance of the three quantization methods will be similar. As discussed and shown in (12), for odd multiples all three re-quantization methods perform the re-quantization in the same manner (and equal input to all three quantizers results in an equal output).

The entropy measure is defined as,

$$E = -\sum_{l=-\infty}^{\infty} p(r_{2,l}) \log_2 p(r_{2,l})$$
(16)

where $r_{2,l}$ is the representation level of quantization bin *l* of the second quantizer. The entropy expressions were derived separately for odd and even values of *k*. For **odd values** of *k*, the expressions for entropy are shown as follows.



Table 1: Entropy of re-quantized data as a function of k for odd values



For even values of k and rounding toward zero, the expressions for the entropy are the following:

Table 2 : Entropy of re-quantized data as a function of k for even values and rounding toward zero



For even values of k with regular rounding, the expressions for the entropy are:

Table 3 : Entropy of re-quantized data as a function of k for even values and regular rounding

The entropy of the re-quantized data is plotted in Figure 7 as a function of k. The first stage quantization step size used is $q_1 = 10$ and the parameter λ is set to 0.1, a value found appropriate for images [6].



Figure 7: Entropy of re-quantized Laplace distribution

It can be observed that for *rounding toward zero*, a much steeper decrease in entropy occurs when k is even. This shows that better compression can be achieved at these points. This can be seen at the sharp transition from 1 to 2 $(q_2=2\cdot q_1)$ and from 3 to 4 $(q_2=4\cdot q_1)$. Whereas, for *regular rounding* a steeper decrease in entropy occurs when k is odd, at the transition from 4 to 5 $(q_2=5\cdot q_1)$ as opposed to the transition from 5 to 6 $(q_2=6\cdot q_1)$. In general, the entropy is substantially higher for *regular rounding*, which means that better compression is achieved using *rounding toward zero*. When comparing to the direct quantization curve, if the reference coarse quantizer is used once on the original data, it can be observed that for the even multiples, the entropy of the re-quantized data is smaller when *rounding toward zero* is used and larger when *regular rounding* is used. At the odd multiples however, the entropy of the re-quantized data is equal to that of the data quantized once with the coarse reference quantizer, as expected. This reinforces the conclusion that deeper compression is achieved at even multiples of the original quantization step size when using *rounding toward zero*.

3.2. Distortion Analysis of a Laplacian Source

The distortion of re-quantized DCT coefficients is considered here, assuming Laplacian distribution before quantization, i.e.,

$$D = \int_{-\infty}^{\infty} (x - \hat{x})^2 f(x) dx = \frac{\lambda}{2} \sum_{l=-\infty}^{\infty} \int_{d_{2,l}}^{d_{2,l+1}} (x - r_{2,l})^2 e^{-\lambda |x|} dx = \lambda \int_{0}^{d_{2,l}} x^2 e^{-\lambda x} dx + \lambda \sum_{l=1}^{\infty} \int_{d_{2,l}}^{d_{2,l+1}} (x - r_{2,l})^2 e^{-\lambda x} dx$$
(17)

using the definition of error caused by quantization. Next, to calculate the error caused by re-quantization, the second stage quantizer decision levels $(d_{2,l})$ and reconstruction levels $(r_{2,l})$ are used. As with the entropy in the previous sub section, the distortion is evaluated as a function of k, separately for odd and even values, for *regular rounding* and for *rounding towards zero*. As before, for odd values, the distortion for both rounding methods and for the direct quantization is equal. Below are derivations for odd values:



Table 4 : Distortion of re-quantized data as a function of k for odd values



For even values of k and for rounding toward zero, the expressions for the distortion are shown below.

Table 5: Distortion of re-quantized data as a function of k for even values and rounding towards zero



For even values of k and for regular rounding, the expressions for the distortion are shown below.

Table 6 : Distortion of re-quantized data as a function of k for even values and regular rounding

The Distortion of the re-quantized DCT coefficients is plotted in Figure 8 as a function of k. As before, the first stage quantization step size used is $q_1 = 10$ and the parameter λ was set to 0.1. The distortion is plotted for both rounding methods as well as for direct quantization, where the data is only quantized once using the reference quantizer.



Figure 8 : Distortion of re-quantized Laplace distribution

It can be observed that for re-quantization using both rounding methods, a much steeper increase in distortion occurs when k is even. This can be seen here at the sharp transition from 1 to 2 $(q_2=2 \cdot q_1)$ and from 3 to 4 $(q_2=4 \cdot q_1)$. In addition, the distortion at the even multiples is the highest for *regular rounding*, second high for *rounding toward zero* and lowest for direct quantization. This means that re-quantization with an even multiple of the original quantization step, additional distortion is inevitable for both rounding methods and it is higher for *regular rounding*. Third, as k increases, the distortion for both re-quantization methods and direct quantization converges to the same value. This can be explained in the following manner. As k increases, the probability weight of the interval

$$\begin{bmatrix} -\frac{q_1}{2}, \frac{q_1}{2} \end{bmatrix} \text{ increases according to}$$

$$P_{zero-bin} = \frac{\lambda}{2} \int_{-kq_1/2}^{kq_1/2} e^{-\lambda|x|} dx = \{q_1 = 10\} = 1 - e^{-5\lambda k}.$$
(18)

Since this interval is quantized to zero in all three methods, when $P_{zero-bin} \approx 1$, the distortion for all three methods converges to the same value. Figure 9 shows $P_{zero-bin}$ as a function of k for $\lambda=0.1$. It can be observed that for

k > 10, $P_{zero-bin} \approx 1$, which explains the similar distortion values in Figure 8, for all three methods. Similar behavior can be observed in Figure 7, for the entropy.



Figure 9 : Probability weight of $P_{zero-bin}$ as a function of k

3.3 Rate-Distortion Analysis

Earlier derivations were used to examine the theoretical rate-distortion behavior of the re-quantized DCT coefficients, assuming initial Laplacian distribution. Figure 10 shows the rate vs. distortion for $q_1 = 10$, $\lambda = 0.1$ and rounding toward zero. It can be seen that there are parts of the curve where moving towards substantially higher distortion does not substantially reduce the rate. For instance, when moving right from $D \approx 85$ to $D \approx 100$ the rate remains $R \approx 1.29$ bit per pixel. This means that there are some re-quantization step sizes that will cause a larger distortion without reducing the rate. Clearly such re-quantization steps are best avoided. This issue is further demonstrated and discussed in the next section.



Figure 10 : Rate- distortion for re-quantization of Laplace PDF

Figure 11 shows the rate-distortion behavior for re-quantization using both *regular rounding* and *rounding toward zero* as well as *direct quantization*. The conclusions remain consistent with previous observations:

- With *rounding toward zero*, the performance is usually superior to *regular rounding* and *direct quantization*, where superior performance means lower distortion for the same rate or lower rate for the same distortion.
- With *regular rounding* there are areas where lowering the rate substantially decreased the distortion only slightly, which is an abnormal behavior for a rate-distortion analysis. The reason is that some requantization steps (even multiples of the original quantization step) perform badly and increase the distortion significantly without improving the compression.
- The points of intersection for all three methods are the odd multiples of the original step size.
- In general, *regular rounding* gives the worst results.



Rate-Distortion for Re-Quantized Laplace PDF

Figure 11: Rate-Distortion for re-quantized Laplacian PDF using three re-quantization methods

4. SIMULATION RESULTS

Two types of quantization were used in the simulation: uniform quantization and quantization using the JPEG quantization matrix. As discussed, two rounding techniques were used: *regular rounding* (where 0.5 is rounded towards 1 and -0.5 is rounded towards -1) and *rounding towards zero* (where 0.5 is rounded towards 0 and -0.5 is rounded towards 0 as well). Figure 12 shows the original "man.tiff" image. Figure 13 shows the "man" image after uniform quantization with q_1 =15. Figures 14-16 show three attempts to re-quantize the image shown in Figure 13.



Figure 12 : "man.tiff"



Figure 13 : $q_1 = 15$, PSNR is 36.7639, rate is 1.3395 bit per pixel, regular rounding

Figure 14 shows this image re-quantized with $q_2 = 29$ and Figure 15 shows the image re-quantized with a slightly coarser re-quantization step of $q_2 = 30 = 2 \cdot q_1$. Rounding towards zero was used for both images.



Figure 14 : $q_2 = 29$, PSNR is 30.5439, rate is 1.1834 bit per pixel, rounding toward zero



Figure 15 q_2 = 30, PSNR is 32.1022, rate is 0.53909 bit per pixel, rounding toward zero

Figure 16 shows the image re-quantized with $q = 30 = 2 \cdot q_1$. This time regular rounding was used.



Figure 16 q_2 = 30, PSNR is 30.0445, rate is 1.1828 bit per pixel, regular rounding

When comparing Figure 14 to Figure 15, it can be easily observed that Figure 15 looks much better than Figure 14. There is less noise and the image appears to be much smoother. The numerical results are summarized in Table 7 below. Even though a coarser re-quantization step was used in Figure 15, the PSNR is higher by approximately 1.5 dB, which is evident visually as well. Moreover, the compression of the image re-quantized with $q_2 = 30$ is much more significant (the required bit-rate is more than two times lower) than the compression of the image re-quantized with $q_2 = 29$. When comparing Figure 15 to Figure 16, the importance of the rounding method becomes obvious.

Those two images are re-quantized with the same step $(q_2 = 30)$ however the rounding method varies. *Rounding toward zero* was used in the image in Figure 15, while *regular rounding* was used in the image in Figure 16. As a result, Figure 16 is much noisier, its PSNR is lower by approximately 2 dB and its bit-rate is more than twice higher.

Fig.13	<i>q</i> ₂ =29	PSNR: 30.54	Rate: 1.18 bits/pixel	Rounding to zero
Fig.14	<i>q</i> ₂ =30	PSNR: 32.1	Rate: 0.54 bits/pixel	Rounding to zero
Fig.15	<i>q</i> ₂ =30	PSNR: 30.04	Rate: 1.18 bits/pixel	Regular rounding

 Table 7 : Numerical results

The following figures demonstrate similar results for quantization using the JPEG quantization matrix. Figure 17 shows the original "girl.bmp" image. Figure 18 shows the same image quantized with the JPEG matrix using the quality factor $Q_I = 0.5$ (which multiplies the matrix). Figures 19 and 20 show results for re-quantization using both rounding methods.



Figure 17 : Original "girl.bmp"



Figure 18 : $Q_1 = 0.5$, *PSNR is 38.619, rate is 1.1157 bit per pixel*



Figure 19 : $Q_1 = 0.5$, $Q_2 = 1$, *PSNR is 33.8481* rate is 0.99316 bit per pixel, regular rounding



Figure 20 : $Q_1 = 0.5$, $Q_2 = 1$, *PSNR is* 35.7367 *rate is* 0.58789 *bit per pixel, rounding toward zero*

In the above figures the same quality factor was used ($Q_2 = 1$). However, *regular rounding* was used in Figure 19 and *rounding toward zero* was used in Figure 20. It can be observed once again that the second image (*rounding toward zero*) is much smoother and less noisy, has a higher PSNR (by almost 2 dB) and has a much lower bit-rate.

The visual results presented above are consistent with the theoretical results of previous sections and show how crucial it is to select the right quantization step. Figure 21 and 22 show results for rate and MSE as a function of the uniform re-quantization step q_2 . The original quantizer step was $q_1=10$. Seven different images (man, camera, army, chip, Lena, gold hill, girl) have been used and the figures show the averaged results. Figure 21 shows the required bit-rate (bits/pixel) of the re-quantized images, relative to the bit-rate required to code the image quantized once with the original quantizer step $q_1=10$. The solid line shows the bit-rate of the reference quantizer (used directly on the original image). The dashed line (with "o" marks) shows the result for re-quantization using *regular rounding*. It can be observed that for this case, the drastic decrease in bit-rate occurs at $q_2=k$ · q_1+1 for even k. The dotted line (with "*" marks) shows the result of re-quantization using *rounding towards zero*. For this case the drastic decrease in the bit-rate occurs on $q_2=k$ · q_1 for even k as shown theoretically in subsection 3.1. Figure 22 shows the MSE as a function of q_2 . The solid line shows the MSE of the reference quantizer (used directly of the original image every time). The dashed line (with "o" marks) shows the result for re-quantization using *regular rounding*. The MSE peaks at $q_2=k$ · q_1 for even k and reaches its local minima at $q_2=k$ · q_1 for odd k, as shown theoretically in subsection 3.2. The dotted line (with "*" marks) shows the result of re-quantization using *rounding towards zero*. The MSE peaks at $q_2=k$ · q_1-1 for even k and also reaches its local minima at $q_2=k$ · q_1 for odd k.

Generally, rounding towards zero gives better results than regular rounding. This is obvious when $q_2=20$ and $q_2=40$. Moreover, the re-quantized images are best compressed when q_2 is an even multiple of q_1 . The MSE is lower when rounding towards zero is used instead of regular rounding.



Figure 21 : *Rate as a function of re-quantization step q*



Figure 22 : MSE as a function of re-quantization step q

Figure 23 shows rate vs. distortion for the three re-quantization methods. The empirical curves obtained below are in full compliance with the behavior and conclusions from the theoretical analysis carried out in Section 3.



Figure 23 : Empirical rate vs. distortion

6. CONCLUSIONS

A re-quantization method for trans-coding JPEG and MPEG intra-frames has been considered and analyzed. The process of re-quantization involves selection of the second stage quantization step, or a quality factor when using the JPEG quantization matrix. The efficiency and performance of the proposed method have been evaluated based on the obtained bit-rate, MSE and visual quality of the re-quantized images. Despite its superior results, the method is straightforward to implement and of low computational complexity. The optimal re-quantization step size can be easily chosen based only on the original quantization step size. The resulting bit-rate is low as well as the MSE.

In this work, we have shown that the bit-rate of the re-quantized image decreases significantly when the requantization step size is an even multiple of the original quantizer step size. This notion was shown for the case of a Laplacian distribution of the DCT coefficients. When the image is re-quantized (with the re-quantization step growing), the probability weights are re-assigned into new bins. The entropy decreases (along with the required bitrate) when the probability weights merge together. The entropy decrease is very significant when even multiples of the original quantization step are used.

On the other hand, the MSE is minimized at odd multiples of the original quantization step size and peaks at even multiples. This has been shown by analyzing the structure of the midtread quantizer used in JPEG and MPEG intra-frames as well as the Laplacian distribution of the DCT coefficients. The distortion of the re-quantized image was compared to the distortion of a reference quantizer that was used directly on the original image. It was shown that the rounding policy of the re-quantization process greatly affects its performance. Rounding quantized DCT values towards zero significantly decreases the MSE for even multiples. In fact, the MSE maximum is then shifted to $q_2=2kq_1-1$. This allows benefiting from the low rate achieved for even multiples, at relatively low distortion. In addition, it was shown that the rounding policy also affects the bit-rate. *Rounding towards zero* achieves the lowest rates at $q_2=2kq_1$ while *regular rounding* shifts the minimum rate point to $q_2=2kq_1+1$.

This work has introduced a rate-distortion model for the re-quantization method, using both rounding methods and direct quantization. Our theoretical analysis supports and explains the experimental results that there are re-quantization step sizes that greatly increase distortion without significantly decreasing the bit-rate, which makes them unsuitable. It has also been shown theoretically that *rounding toward zero* always outperforms *regular rounding*. Our conclusion is that most efficient trans-coding is achieved by selecting a re-quantization step that is an even multiple of the original quantization step and rounding the re-quantized coefficients towards zero. These results could be helpful in improving currently available trans-coding systems, especially for real-time applications.

6. REFERENCES

[1] I. Ahmad, X. Wei, Y. Sun and Y.Q. Zhang, "Video Transcoding: An Overview of Various Techniques and Research Issues", *IEEE Transactions on Multimedia*, Vol. 7, NO. 5, pp. 793-804, Oct. 2005.

[2] A. Pedro, A. Assuncao and M. Ghanbari, "A Frequency-Domain Video Transcoder for Dynamic Bit-Rate Reduction of MPEG-2 Bit Streams", *IEEE Transactions on Circuits And Systems For Video Technology*, Vol. 8, NO. 8, pp. 953 - 967, Dec. 1998

[3] A. Leventer and M. Porat, "Towards optimal bit-rate control in video transcoding", *ICIP - International Conference on Image Processing*, Vol. 3, pp.265-8, Sep. 2003.

[4] A. Leventer and M. Porat, "On Bit Allocation in Video Coding and Transcoding", *International Conference on Visual Information Engineering*, pp. 254 – 257, Jul. 2003.

[5] O. Werner, "Requantization for transcoding of MPEG-2 Intraframes," *IEEE Transactions on Image Processing*, Vol. 8, No. 2, pp.179-191, Feb. 1999.

[6] S. R. Smoot, L. A. Rowe, and E.F. Roberts, "Laplacian model for AC DCT terms in image and video coding," *Proceedings of the 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, Monterey CA, pp.60-9, 2003.

[7] Z. Fuhao, L. Zhengding and L. Hefei, "Statistical Model of Quantized DCT Coefficients", *ICSP* – 7th *International Conference on Signal Processing*, Vol. 3 pp. 2572-2575, 2004.

[8] B. Shen, "Modeled Analysis on Requantization Error", *IEEE 7th Workshop on Multimedia Signal Processing*, pp. 1-4, Oct. 2005.

[9] N. Hait, "Model-Based Transrating of Coded Video", Technion's Research Thesis, pp. 14-16 Nov. 2007.

[10] H.H Bauschke, C.H.Hamilton, M.S Macklem, J.S McMichael, and N.R Swart, "Recompression of JPEG Images by Requantization", *IEEE Transactions on Image Processing*, Vol. 12, Issue 7, pp. 843 - 849, July 2003