Quality-Controlled Compression of Sets of Images (extended abstract)

Viresh Ratnakar
University of Wisconsin-Madison
Computer Sciences Department
Madison, WI 53706
Phone: (608) 262-6627
Email: ratnakar@cs.wisc.edu

Miron Livny
University of Wisconsin-Madison
Computer Sciences Department
Madison, WI 53706
Phone: (608) 262-0856
Email: miron@cs.wisc.edu

Abstract

In this paper, we propose the QCLICS algorithm for quality-controlled JPEG compression of image sets. Lossy compression of images results in a tradeoff between compression and quality: the greater the compression, the poorer the reulting image quality. Quality-controlled compression has two goals: meeting the desired compression/quality specification accurately, and optimizing the compression-quality tradeoff. When a large set of images is to be compressed into a limited amount of disk space, it is desirable to optimize the compression-quality tradeoff for individual images as well as across images. This is an important issue in multimedia CD-ROM production, where hundreds or thousands of digital images need to be compressed to fit in the available space. The QCLICS algorithm uses the RD-OPT algorithm [RL95, RL96] for optimizing JPEG quantization tables for individual images, and uses a dynamic programming strategy to apportion the target compression/quality to the individual images. The complexity overhead of optimizing across the images is minimal, while significant gains in compression/quality are obtained. QCLICS offers flexibility in choosing the image quality metric, and allows various constraints to be placed on compression/quality for individual images. We present performance results for several sets of images. These images were taken from various scientific sources at the University of Wisconsin-Madison, such as microscopy images of cells, crystals, brains, and spacecraft images of Venus.

1 Introduction

Digital images occupy a major chunk of today's computer media. On account of their high space and bandwidth requirements, digital images are typically stored and transmitted in compressed form. Moreover, since lossless compression does not reduce image size significantly (at most by a factor of 2, typically), lossy compression is used. With lossy compression, any amount of size reduction is possible, albeit with an associated loss in quality. The greater the compression, the poorer the quality of the compressed image with respect to the original. This compression-quality tradeoff depends upon specific image characteristics as well as the compression technique used. Optimizing this tradeoff for individual images has been the subject of much previous work [RL95, RL96, MS93, WG93, CR95]. We use the term *quality-controlled compression* to refer to compression where user-specified compression/quality objectives are met accurately while the compression-quality tradeoff is optimized.

In this paper, we address the problem of quality-controlled compression of large sets of images. A set of images is simply a collection of images grouped together for some reason. For example, a collection of satellite images shot periodically, or simply a set of arbitrary images to be put on a single CD-ROM. In addition to optimizing compression-quality tradeoffs for individual images in a set, it is important to allocate compression-quality quotas optimally to the individual images out of a fixed total chunk. An example application for this is multimedia CD-ROM production, where the goal would be to find, for instance, the best way to pack a set of 1000 images in the available disk space of 20 Megabytes. Several constraints, such as limits on worst acceptable individual image quality or size, may also be placed. For example, a multimedia archive delivering images across a network may be asked to provide Quality-of-Service guarantees, such as transmitting a certain number of images per second across a certain bandwidth. This would translate to minimum limits on compression for individual images.

Several lossy image compression techniques are in common use, the most popular one being the JPEG standard [PM93]. JPEG uses the Discrete Cosine Transform (DCT) [ANR74] to transform the image data to a spatial frequency domain. Each spatial frequency coefficient is stored with a certain degree of accuracy, which can be varied to vary the extent of compression or quality. We have developed an algorithm (called the RD-OPT algorithm) [RL95, RL96] for quality-controlled JPEG compression. RD-OPT uses image-specific DCT coefficient distributions to predict and optimize the compression-quality tradeoff. Here we use and extend RD-OPT to optimize JPEG compression of sets of images. The resulting algorithm is referred to as the QCLICS (Quality-Controlled Lossy Image Compression for Sets) algorithm. Specifically, QCLICS addresses the problem of maximizing the "quality" of a set of images while keeping the compressed size under a given target, or minimizing the compressed size while keeping the quality above a certain target.

2 Compression and Quality for a Set of Images

In this section, we fix some notation and address the issue of measuring compression and quality for image sets.

We assume 8-bit precision for pixel intensities here, but the results are valid for any precision. For a $W \times H$ grayscale image I, the number of pixels P is WH, and the uncompressed image size is P bytes. For color images, JPEG compression usually involves a preprocessing step in which the image is transformed from (R,G,B)-valued pixels to (Y,Cb,Cr)-valued pixels, possibly followed by subsampling in the chrominance planes Cb and Cr [PM93]. Let the sum of the number of pixels in each plane (Y, Cb, and Cr) after preprocessing be P. The uncompressed image size (after preprocessing) is P bytes.

Compression is usually measured as *rate*, which is defined as the number of bits used per pixel in the compressed image. Thus, if the size of the compressed image is S bytes, then the rate R is equal to $8S/P$ bits per pixel (bpp).

Measuring the quality of a compressed image is the subject of much debate, as there is no metric that adequately captures the amount of loss the human eye can detect. Further, specific applications may require their own customized metrics. Most of the commonly used measurements, however, are based on *distortion*. The distortion D of a compressed image is the mean squared error between pixel values in the compressed image and the original image. Distortion may also be measured in the frequency domain, with different weights attached to different frequencies. Further, for color images, different weights may be attached to distortions in different planes. All these weighted variants of distortion can be used in QCLICS . Peak Signal to Noise Ratio (PSNR) is the most commonly used quality

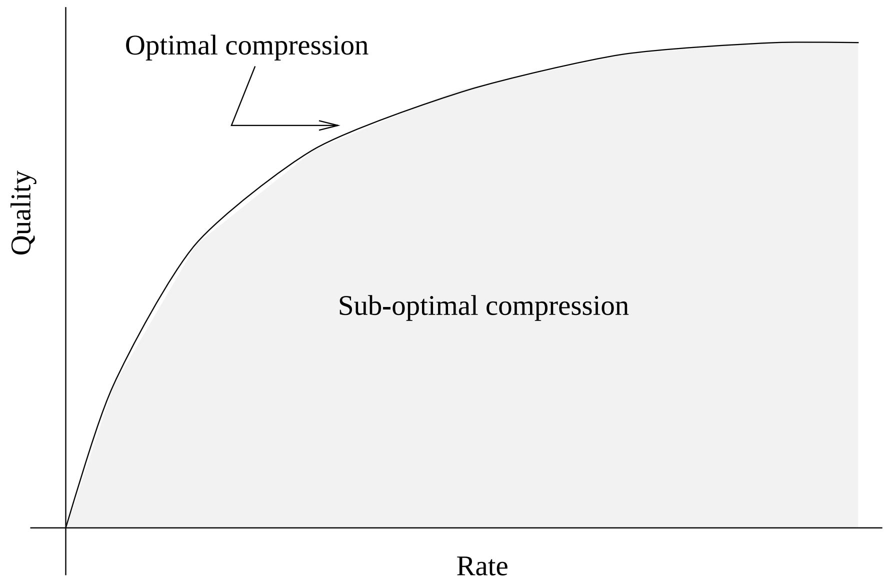


Figure 1: Rate-quality tradeoff in lossy compression

measurement based on distortion. The quality Q of the compressed image, measured as PSNR, is given by

$$Q = 10 \log_{10} \frac{255^2}{D} \text{ dB}.$$

With most lossy compression techniques, it is possible to vary certain parameters to achieve different values of R and Q. For a given value of R, it is desirable to achieve the best possible Q. With sets of images too, it would be desirable that the rate-quality tradeoff be optimized for each individual image. The rate-quality curve (Figure 1) for an image depends on several image characteristics (such as presence of high frequencies, sharp edges, etc.) along with the compression technique used. Naive choices of compression parameters may result in sub-optimal rate-quality tradeoffs [RL94]. For JPEG, the compression parameters are the quantization table and the thresholding strategy [PM93, RL96, CR95]. We use the RD-OPT algorithm to achieve near-optimal tradeoffs.

2.1 Image sets

Let $\mathcal{I} = \{I_1, I_2, \ldots, I_N\}$ be a set of N images, with P_i being the total number of pixels in the i^{th} image. Let S_i (R_i) denote the size (rate) and Q_i the quality of the compressed images.

We can measure compression for the set $\mathcal{I}$ in two ways, averaging either the rates or the sizes. These two will be different if the number of pixels in different images is different. Averaging rate assigns equal importance to each pixel in each image, while averaging size assigns equal importance to each image, irrespective of its number of pixels. Assuming that the set $\mathcal{I}$ is not a collection of images grouped together arbitrarily, it should be possible to determine which average to use. For example, if the set $\mathcal{I}$ shows the same picture at many different resolutions, average rate should be used; on the other hand, if $\mathcal{I}$ is a collection of facial pictures with minor differences in width and height, average size should be used. In general, the *rate* R of the compressed set $\mathcal{I}$ is given by

$$R = \sum_{i=1}^{N} r_i,$$

where $r_i = R_i w_i^r$, w_i^r being a non-negative weight assigned to the i^{th} image. We will use average size for our examples, for which,

$$w_i^r = P_i / P,$$

where $P = \sum_{j=1}^{N} P_j$. In this case, the total size of the compressed set $\mathcal{I}$ is $RP/8$ bytes.

Another possibility for measuring compression is to take the maximum of all R_i (or S_i). Keeping this maximum under a target while maximizing quality, however, reduces simply to choosing equal rates (or sizes) for all images. We allow *constraints* in the form of bounds on r_i, along with a target value for $R = \sum r_i$. This general case includes the maximum-rate measurement of compression.

Quality of a compressed set of images can also be measured by averaging the individual qualities. Again, different weights may be assigned to different images.

$$Q = \sum_{i=1}^{N} q_i,$$

where $q_i = Q_i w_i^q$, w_i^q being a non-negative weight assigned to the i^{th} image. For our examples, we use the simple average, in which case, $w_i^q = 1/N$.

Here too, the minimum quality can be used as the measurement, but keeping this minimum above a given target while maximizing compression reduces to choosing equal qualities for all images. We allow constraints on q_i too, which captures this quality measurement.

3 Rate-Quality Optimization for Sets of Images

Given a set $\mathcal{I}$ of N images, non-negative weights w_i^r and w_i^q, quality constraints q_i^L, q_i^H, and rate constraints r_i^L, r_i^H, the optimization problem can be phrased in two ways.

1. Given a target rate R^*, compress the image set $\mathcal{I}$ such that $R \leq R^*$ and Q is maximized.

2. Given a target quality Q^*, compress the image set $\mathcal{I}$ such that $Q \geq Q^*$ and R is minimized.

In both cases, the individual compression rates ($r_i = R_i w_i^r$) and qualities ($q_i = Q_i w_i^q$) must satisfy the constraints:

1. $r_i^L \leq r_i \leq r_i^H$, and

2. $q_i^L \leq q_i \leq q_i^H$, for each $i = 1, \ldots, N$.

3.1 The QCLICS Algorithm

The QCLICS algorithm has three steps:

1. RD-OPT.

2. Dynamic programming.

3. Compression.

We use the RD-OPT algorithm to get a set of constraint-satisfying rate-quality points for each image, and then use a dynamic programming approach to combine these. This results in the selection of compression parameters for each image, such that the rate/quality objective is met, and the rate-quality tradeoff is nearly optimal. Finally, the images are compressed according to their selected parameters. We now present each of these steps in detail.

The RD-OPT step

Run RD-OPT on each image I_i in $\mathcal{I}$. Get n rate-quality points $\{(r_i^j, q_i^j) \mid j = 1, \ldots, n\}$ and the associated quantization tables $\{T_i^j \mid j = 1, \ldots, n\}$, such that the constraints

$$r_i^L \leq r_i^j \leq r_i^H, \text{ and, } q_i^L \leq q_i^j \leq q_i^H$$

are satisfied at each point, and the rates r_i^j are spread uniformly over the range enforced by the constraints. Here the number of points, n, is a user-supplied value. Higher values of n result in finer granularity of the search space, but increase the time complexity of the dynamic programming step.

The dynamic programming step

Discretize each r_i^j as $\rho_i^j = \mathit{RoundOff}(r_i^j \times B)$, where B is a large integer. Let ρ^M be the discretized maximum total rate for $\mathcal{I}$. If a target rate R^* for $\mathcal{I}$ has been specified then $\rho^M = \mathit{RoundOff}(R^* \times B)$, while if a target quality has been specified then $\rho^M = \sum_{i=1}^{N} \rho_i^n$.

Let BestQ be a table with N rows and ρ^M columns. The entries in BestQ have the following interpretation: $\mathit{BestQ}(i, \rho)$ is the highest sum of qualities for images I_1 through I_i such that the sum of their discretized rates is exactly ρ. Initialize BestQ to be $-\infty$ everywhere.

The table BestQ is easily filled, starting from row 1, and marching down the rows, using the information about the i^{th} image and the values in the $(i-1)^{\text{th}}$ row to fill the i^{th} row:

$$\mathit{BestQ}(i, \rho) = \max \left\{ q_i^j + \mathit{BestQ}(i-1, \rho') \;\middle|\; \begin{array}{l} 1 \le j \le n \\ \rho' \ge 0 \\ \rho' + \rho_i^j = \rho \end{array} \right\}.$$

The choice of j made for each $\mathit{BestQ}(i, \rho)$ is recorded in the $N \times \rho^M$ array *Choice*.

1. Initialize each $\mathit{BestQ}(i, \rho)$ to $-\infty$.

2. Initialize row 1:

for $j := 1$ **to** n

if $q_1^j > BestQ(1, \rho_1^j)$ **then**

$BestQ(1, \rho_1^j) := q_1^j$

$Choice(1, \rho_1^j) := j$

3. Fill remaining rows:

for $i := 2$ **to** N

for $j := 1$ **to** n

for $\rho' := 1$ **to** $\rho^M - \rho_i^j$

if $q_i^j + BestQ(i-1, \rho') > BestQ(i, \rho_i^j + \rho')$ **then**

$BestQ(i, \rho_i^j + \rho') := q_i^j + BestQ(i-1, \rho')$

$Choice(i, \rho_i^j + \rho') := j$

Now, if a target rate R^* has been specified, find the highest value of ρ such that $\rho \leq RoundOff(R^* \times B)$ and $BestQ(N, \rho) > -\infty$. If a target quality Q^* has been specified, then find the least ρ such that $BestQ(N, \rho) \geq Q^*$. From this starting point in the N^{th} row, the choice j_i^* for each image i can be recovered as follows:

for $i := N$ **down to** 1

$j_i^* := Choice(i, \rho)$

$\rho := \rho - \rho_i^{j_i^*}$

The compression step

Compress each image using the selected DCT quantization tables. Thus, I_i is compressed using the tables $T_i^{j_i^*}$. The tables may be simply the quantization tables produced by the

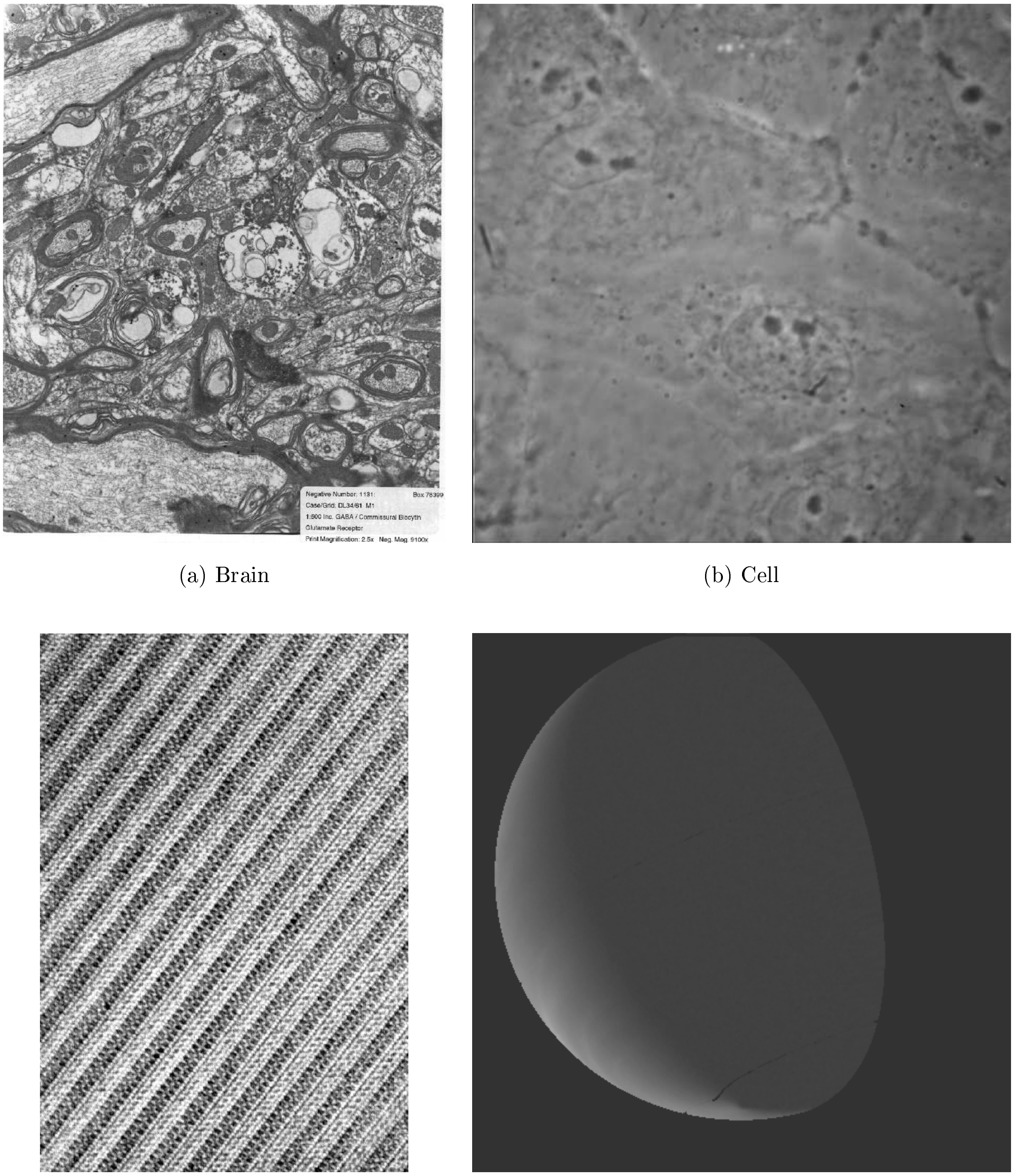
RD-OPT algorithm in [RL95], or the quantization and thresholding tables generated by the extended RD-OPT algorithm of [RL96]. In the latter case, the rate-quality tradeoff will be closer to the optimum. In either case, the resulting rate and quality will be very close to the predictions.

4 Performance

We present performance results for four sets of images. These sets are:

1. Brain: 11 grayscale images of cat's brains.
2. Cell: 31 grayscale images of a cell.
3. Crystal: 14 grayscale images of various crystals.
4. Venus: 17 grayscale images of Venus.

A sample image from each set is shown in Figure 2. For each set, we ran QCLICS with several constraints and objectives. Average sizes were used for rate (that is, each $w_i^r = P_i/P$), and simple averaging was done for quality (that is, each $w_i^q = 1/N$). When a target rate R^* was specified, each R_i was constrained to be in the range $(R^* - \delta_r, R^* + \delta_r)$, and when a target quality Q^* was specified, each Q_i was constrained to be in the range $(Q^* - \delta_q, Q^* + \delta_q)$, for various values of δ_r and δ_q. The number of rate-quality points used per image (n) was 50, and the discretization constant B was 5000. The rates predicted by RD-OPT were corrected by compressing the image once at a median rate and observing the error. JPEG compression was done using the Independent JPEG Group's `cjpeg` software.



(a) Brain

(b) Cell

(c) Crystal

(d) Venus

Figure 2: Sample images

Name	R^* (bpp)	δ_r (bpp)	Predicted R (bpp)	Actual R (bpp)	Predicted Q (dB)	Actual Q (dB)
Brain	1.20	0.30	1.20	1.21341	29.623	29.637
Cell	0.75	0.25	0.75	0.75119	49.796	49.211
Crystal	1.50	0.50	1.50	1.41536	35.733	35.762
Venus	0.75	0.25	0.75	0.73924	57.993	59.342

Figure 3: QCLICS performance with target R^*

Name	Q^* (dB)	δ_q (dB)	Predicted R (bpp)	Actual R (bpp)	Predicted Q (dB)	Actual Q (dB)
Brain	37.0	04.0	1.79300	1.86301	33.000	33.015
Cell	50.0	04.0	0.83480	0.79224	50.001	49.424
Crystal	35.0	10.0	1.08140	1.01930	35.002	35.017
Venus	55.0	05.0	0.32920	0.37846	55.002	55.528

Figure 4: QCLICS performance with target Q^*

Figure 3 shows the performance of QCLICS when R^* was specified, and Figure 4 shows the performance of QCLICS when Q^* was specified. The accuracy of predicted rates and qualities depends upon the accuracy of RD-OPT's predictions. The quality predictions are highly accurate (rms error is predicted with an accuracy of ± 0.25), while the rate predictions are typically within ± 0.05 bits per pixel [RL95]. All constraints and targets are satisfied within these tolerances. Note that after the dynamic programming step, *any* target rate/quality is readily obtained for the given constraints. Thus, if higher accuracy is desired, the compression step can be repeated after adjusting the target and selecting a new set of quantization tables $T_i^{j_i^*}$.

Figure 5 shows the average running time per image for each step of QCLICS for each image set, on an IBM POWERstation 370. The running time is dominated by the time taken in the RD-OPT step. The above results used the thresholding strategy of [RL96]. For greater speed, more limited or no thresholding may be used. Further, the range of values in the quantization tables may be reduced. For example, on the Brain set, the average RD-

Name	RD-OPT	Dynamic programming	Compression	Total
Brain	33.79	0.19	1.70	35.68
Cell	07.63	0.16	0.69	08.48
Crystal	19.34	0.22	1.81	21.37
Venus	11.75	0.08	0.56	12.39

Figure 5: QCLICS average running time per image in seconds

OPT time per image can be reduced to 25.5 seconds by not using thresholding. Of course, these reductions in running time move the rate-quality points farther away from the optimal tradeoff. The time spent in the dynamic programming step of QCLICS is linear in n and B. However, this is not significant when compared with the RD-OPT time.

References

[ANR74] Ahmed, N., Natarajan, T., and Rao, K. R. Discrete Cosine Transform. *IEEE Trans. Computers*, C-2390-3, Jan. 1974.

[CR95] Crouse, M. and Ramchandran, K. JPEG optimization using an entropy-constrained quantization framework. *Proceedings of Data Compression Conference*, pages 342–351, 1995.

[MS93] Monro, D. M. and Sherlock, B. G. Optimum DCT Quantization. *Proceedings of Data Compression Conference*, pages 188–194, 1993.

[PM93] Pennebaker, W. B. and Mitchell, J. L. *JPEG Still Image Data Compression Standard*. Van Nostrand Reinhold, New York, 1993.

[RL94] Ratnakar, V. and Livny, M. Performance of Customized DCT Quantization Tables on Scientific Data. *Science Information Management and Data Compression Workshop Proceedings, NASA Conference Publication 3277*, pages 1–8, Sept 1994.

[RL95] Ratnakar, V. and Livny, M. RD-OPT: An Efficient Algorithm For Optimizing DCT Qunatization Tables. *Proceedings of Data Compression Conference (Also, Technical Report 1257, Dept of Computer Sciences, UW-Madison)*, pages 332–341, 1995.

[RL96] Ratnakar, V. and Livny, M. Extending RD-OPT with Global Thresholding for JPEG Optimization. *Proceedings of Data Compression Conference*, 1996.

[WG93] Wu, S. and Gersho, A. Rate-constrained picture-adaptive quantization for JPEG baseline coders. *Proc. Inter. Conf. Acoustics, Speech and Signal Processing*, 5:389—392, April 1993.