/9-95 ... /' 9

N95- 14586

p. 8

Performance of Customized DCT Quantization Tables on Scientific Data*

Viresh Ratnakar
University of Wisconsin-Madison
Computer Science Department
Madison, WI

Miron Livny
University of Wisconsin-Madison
Computer Science Department
Madison, WI

Abstract

We show that it is desirable to use data-specific or *customized* quantization tables for scaling the spatial frequency coefficients obtained using the Discrete Cosine Transform (DCT). DCT is widely used for image and video compression [MP89, PM93] but applications typically use default quantization matrices. Using actual scientific data gathered from divers sources such as spacecrafts and electron-microscopes, we show that the default compression/quality tradeoffs can be significantly improved upon by using customized tables. We also show that significant improvements are possible for the standard test images *Lena* and *Baboon*. This work is part of an effort to develop a practical scheme for optimizing quantization matrices for any given image or video stream, under any given quality or compression constraints.

1 Introduction

We are developing an environment for "production-mode" compression of still-image and video data, where the user can specify constraints on the desired quality and compression ratio, and the compressor produces the best results under those constraints without any human assistance. Under both the JPEG and MPEG compression standards, quality and compression-ratio can be varied by varying the DCT coefficients' quantization table. Most existing encoders simply use a default table and scale it up or down by a small factor to achieve different qualities/compression-ratios. This paper shows that customized quantization tables can outperform scaled default tables to a high degree. We are exploring efficient algorithms for designing these customized tables.

The test images and video streams used for the performance study were some spacecraft images (*Earth*, *Venus*), some molecular-biology images (*Cell*, *Egg*), and some standard images

*This work was partially supported by NASA grant NAGW-3914 and NSF grant IRI-9224741.

used in image compression literature (*Lena*, *Baboon*). In every case, substantial gains were obtained. At a given bit rate, Peak Signal-to-Noise Ratio (PSNR) could usually be improved by about 1-2 dB, while at a given PSNR, bit rate could be reduced by about 0.2 bits per pixel (bpp).

The rest of this paper is organized as follows: Section 2 outlines the use of DCT for image compression and the role of quantization. Section 3 presents the performance of customized quantization tables. Subsection 3.1 shows the results for two standard images, *Lena*, and *Baboon*. Subsection 3.2 shows the results for four scientific data streams. Conclusions are presented in section 4.

2 Discrete Cosine Transform

JPEG and MPEG work by dividing each image (or frame) into blocks of size 8×8 and transforming each block using the Discrete Cosine Transform (DCT). This transformation results in an 8×8 block F of 64 coefficients for each image block f. These coefficients define the unique representation of f as a linear combination of 64 predefined *basis blocks* of the DCT.

The basis blocks capture different spatial frequencies of an image. $F(0,0)$ is the coefficient of the term with zero spatial frequency (the DC component) while $F(7,7)$ is the coefficient of the term with highest frequencies in the x and y directions. Details can be found in [RY90].

The advantage in using DCT is that it compacts most of the signal energy into the low frequency coefficients. The human eye is relatively insensitive to high spatial frequencies which can hence be ignored. In fact, the entire block F of coefficients is quantized using some 8×8 quantization table Q. Thus, the value $F(u,v)$ is stored as the integer closest to $\frac{F(u,v)}{Q(u,v)}$. Higher frequencies are quantized more coarsely (i.e. with a greater value of $Q(u,v)$) than lower frequencies. A good proportion of the high-frequency coefficients get quantized to zero. This enables the block to be compressed efficiently using entropy coding techniques such as Huffman coding or arithmetic coding [Jai89, PM93]. However, this compression is lossy as the reproduced values of the coefficients will not necessarily be the same as the original values.

The quantization table Q ultimately determines the compression ratio and the quality of the reconstructed image. JPEG allows only a fixed table for the entire image. Most encoders set this table to the table suggested in the standard [PM93]. For better quality, the entire table is scaled down while for higher compression, the entire table is scaled up by a small factor which is called the *qscale*.

MPEG allows the quantization table to be changed along a video stream. In addition, the table used may be scaled up or down by multiplying by a *qscale* on a per-macroblock basis. This scaling is done by heuristically determining regions of low activity and high activity and adjusting *qscale* accordingly (See, for example, [CP84]). The table used is generally the one suggested in the standard [MP89].

3 Performance of customized quantization tables

We used Peak Signal-to-Noise Ratio as the measure of quality of a decompressed image. If I is an $M \times N$ grayscale image with pixel values in the range [0..255], and I is approximated by image I', then

$$\text{PSNR} = 10.0 * \log_{10}\left(\frac{\sum_{i,j}(255)^2}{\sum_{i,j}(I(i,j) - I'(i,j))^2}\right).$$

An approximation with PSNR greater than about 37.0 dB is usually indistinguishable from the original image, to the human eye.

Degree of compression was measured as bits per pixel (bpp) used. For the 8-bit grayscale images used, compression ratio is equal to (8/bpp): 1.

For each test image or video stream, we plotted PSNR vs bpp for customized tables and default tables. To obtain customized tables, for every image/stream at every PSNR we searched a wide range of quantization tables to find the best performance (in terms of actual bit rate), using trial and error. The default tables were obtained by multiplying the standard tables suggested by JPEG (for still images) and MPEG (for video streams) by a *qscale* in the range $\frac{1}{8} \cdots \frac{31}{8}$, as is allowed under the two standards [MP89, PM93].

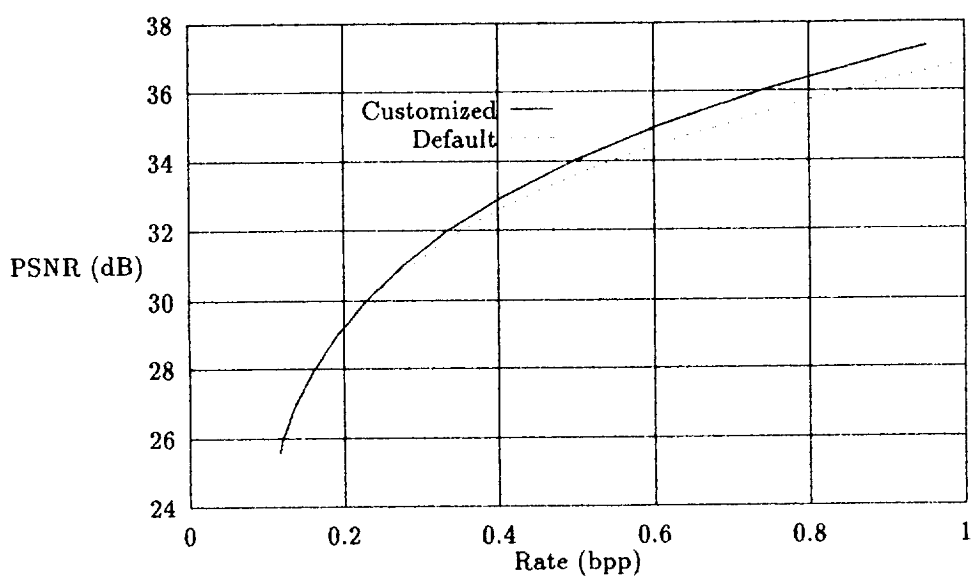

Figure 1: Quality vs Rate curves for *Lena*

The range of the plots was chosen as 0..1 bpp. The lowest bit rate plotted for the default curve was the one achieved by multiplying the default table by a *qscale* of 31/8. For the customized curve, the lowest bit rate was that achieved by maximizing all table entries. The highest bit rate for the default curve (*qscale* = 1/8) was typically in the 0.9 to 1.2 bpp range. All the plots go up to 1 bpp for ease of comparison. The images are not reproduced here because of space constraints.

3.1 Performance results for *Lena* and *Baboon*

Figures 1 and 2 show the results for the standard 512 × 512 8-bit grayscale images *Lena* and *Baboon*, respectively. The default table used was that suggested by the JPEG standard [PM93]. The advantage of using customized tables is seen to be more at higher rates and better qualities. For example, for *Lena*, an improvement of about 0.5 dB in quality can be obtained at rate of 0.8 bpp. The improvement is more pronounced for *Baboon* which

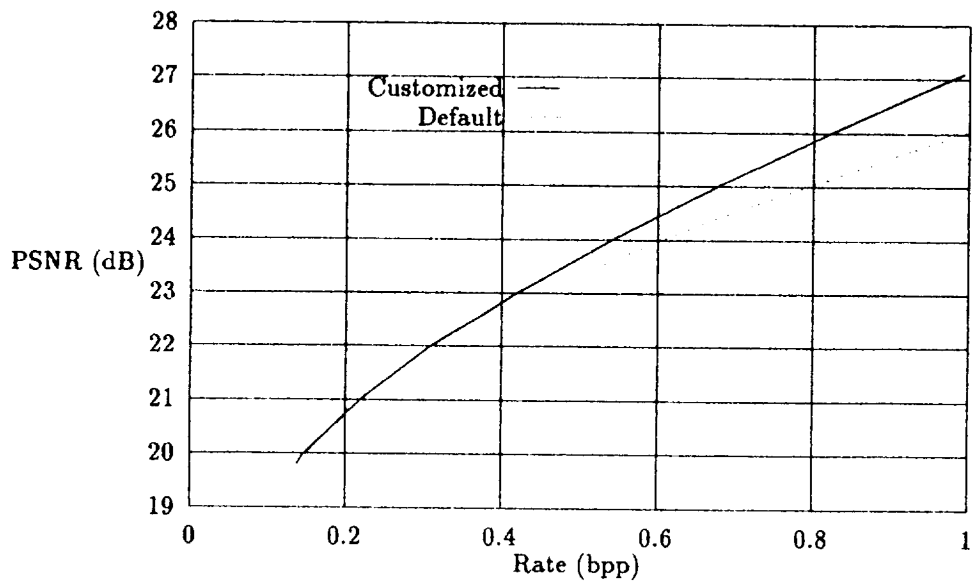

Figure 2: Quality vs Rate curves for *Baboon*

offers a gain of about 1 dB at 0.8 bpp. Comparing bit rates at fixed qualities, we see that a reduction by about 0.1 bpp is achieved at 35 dB for *Lena*. Once again, the difference is more pronounced for *Baboon* where, for example, a reduction of nearly 0.2 bpp can be seen at about 26 dB. It is interesting to note that *Baboon* has a lot of high frequency content and is known to be hard to compress. In general, the improvements offered by customized tables were greater for hard-to-compress images. For such images, the default tables were not able to quantize the high frequencies efficiently, while the customized tables did a much better job. This was seen from the fact that the variances of quantized high-frequency coefficients were lower with customized tables than with default tables.

3.2 Performance results for scientific data

All the results in this section are based on video streams compressed using the I-frames of MPEG-I. Figure 3 shows the results for a stream of 320 × 300 pictures of Earth taken from a satellite. The original pictures were in raw 8-bit grayscale format. Figure 4 shows the results for a stream of 480 × 480 pictures of Venus shot by a NASA spacecraft. Again, the original pictures were in raw 8-bit grayscale format. Figure 5 plots the results for a stream of 352×240 8-bit grayscale pictures of a cell. Finally, Figure 6 refers to a computer generated sequence of 320×240 8-bit grayscale images. For *Earth*, quality improvements went up

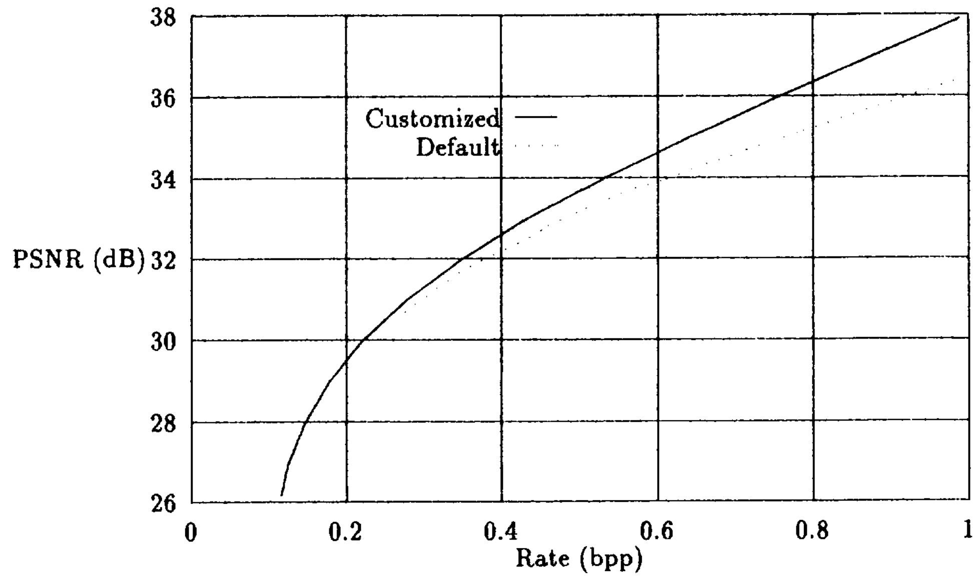

Figure 3: Quality vs Rate curves for *Earth*

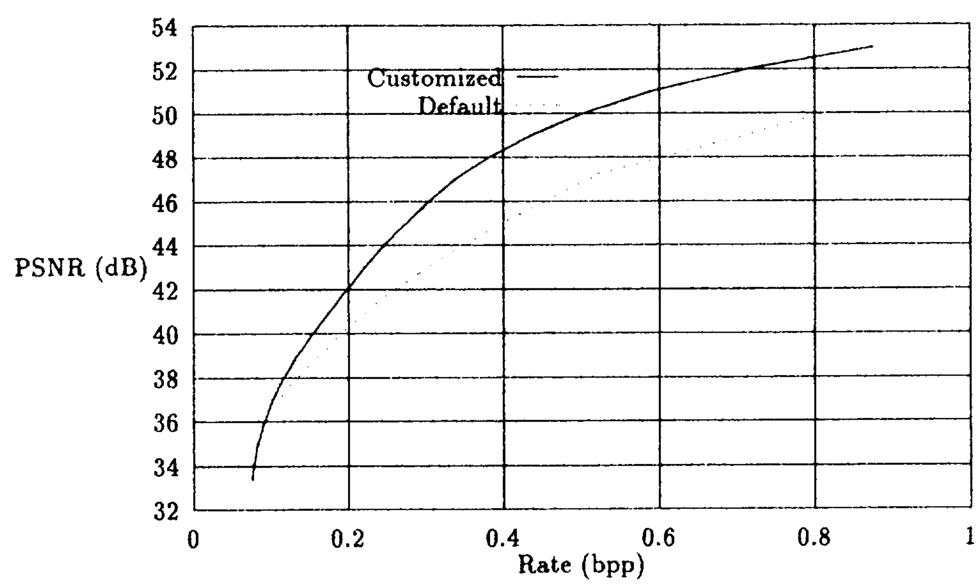

Figure 4: Quality vs Rate curves for *Venus*

to about 2 dB while bit rate reductions up to 0.2 bpp were obtained. *Venus* was rather easy

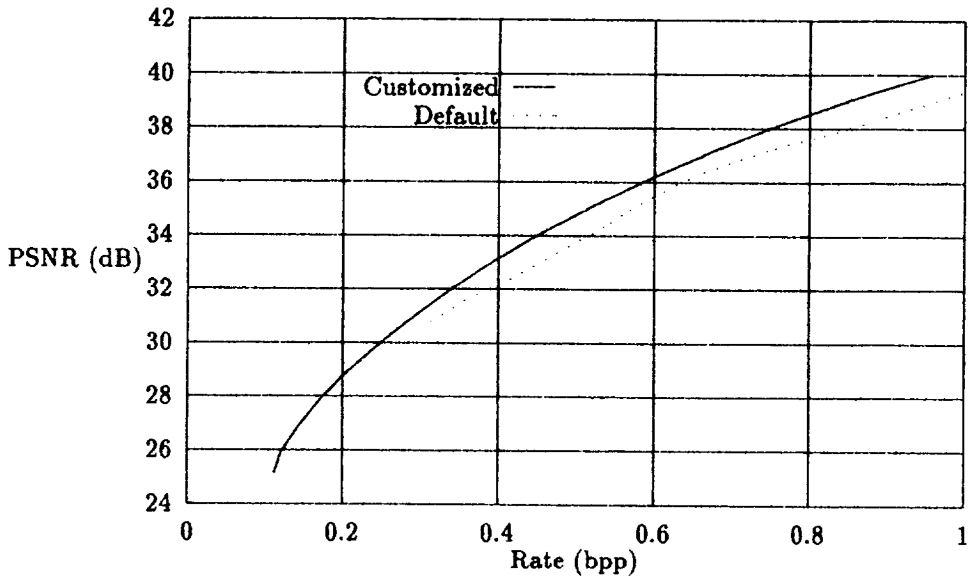

Figure 5: Quality vs Rate curves for *Cell*

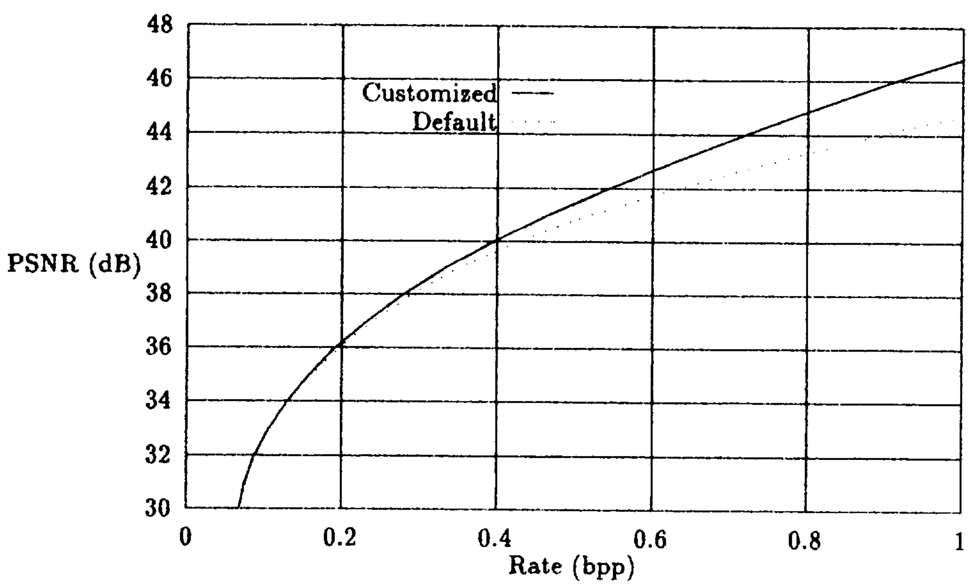

Figure 6: Quality vs Rate curves for *Egg*

to compress as can be seen from the fact a PSNR of 42 dB is achievable at merely 0.2 bpp. The plot for *Venus* shows that the improvement in quality varied between 2 dB (at 0.2 bpp) and 3 dB (at 0.6 bpp). The reduction in rate varied between 0.08 bpp (at 42 dB) and 0.3 bpp (at 50 dB).

For *Cell*, gains in quality were about 1 dB at every bit rate, while reduction in bit rate varied between 0.05 (at 32 dB) to about 0.1 bpp (at 39 dB). The last stream, *Egg*, displayed gains in PSNR up to 2 dB (at 0.9 bpp), and rate reduction up to 0.2 bpp (at 45 dB). The PSNR for *Egg* was better than that for *Cell* at every bit rate. This is to be expected as the stream *Egg* was generated using computer animation and had lesser high-frequency content.

We can see that customized quantization tables improved quality and compression-ratio for every image and video stream. The improvements varied a bit in amount across different images and streams, but were usually substantial enough to justify the use of customized tables, especially at bit rates exceeding 0.6 bpp.

4 Conclusion

Using image data gathered from widely different sources, we have shown that the performance of default tables can always be significantly improved upon. A reduction of 0.2 bpp for 1000 pictures of earth, each 320×300 8-bit grayscale, translates to an additional saving of around 2.4 Megabytes.

We are developing algorithms to design customized quantization tables efficiently, to exploit these possible savings in bit rate and gains in quality. A good choice of the quantization table Q becomes extremely important for production-mode compression environments. In production-mode, the compressor might be presented with widely varying image- and stream-types. A naive choice of Q might give poor performance, as we have seen (particularly for images with large high-frequency content).

For both the default and customized cases we have shown the quality/compression tradeoffs. But deciding which point on the curve to choose, given some constraints (such as exact values or ranges of tolerance for rate and quality), is also a non-trivial problem.

We also tried to exploit customized tables further by adaptively scaling *qscale* on a per-macroblock basis. This did not yield any improvement in PSNR in most cases. However, adaptive scaling does offer better *visual* quality. Further work is needed to detect and exploit scene changes. A new customized table should be introduced on a scene change. Further gains can also be obtained by similarly customizing quantization tables for the non-intracoded frames of MPEG with motion compensation.

Acknowledgements

We thank researchers in the departments of Space Sciences and Molecular Biology at University of Wisconsin-Madison for providing us with video data.

References

[CP84] Chen, W. and Pratt, W. K. Scene adaptive coder. *IEEE Trans. Comm.*, 32:225–232, 1984.

[Jai89] Jain, A. K. *Fundamentals of Digital Image Processing.* Prentice Hall, Englewood Cliffs, NJ, 1989.

[MP89] MPEG proposal package description, July 1989. Document ISO/WG8/MPEG/89-128.

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

[RY90] Rao, K. R. and Yip, P. *Discrete Cosine Transform: Algorithms, Advantages, Applications.* Academic Press, Inc, San Diego, California, 1990.