FREQUENCY CONTENT ORIENTED MODIFICATION OF QUANTIZATION MATRIX IN DCT – BASED COMPRESSION

Daniela Tarniceriu, Valeriu Munteanu, Florin Beldianu

Technical University "Gh. Asachi", Faculty of Electronics and Telecommunications, Blvd. Carol I, No. 11, Iasi, 700506

Abstract: In order to improve the compression rate for comparable values of error metrics, we propose a method to modify the standard quantization matrix used in JPEG, according to image frequency content. Unlike the method which allows varying levels of image compression and quality, by scaling all the elements in the quantization matrix by the same factor, we propose a way to modify differently the quantization step for low and high frequencies, respectively.

Keywords: Image compression, quantization matrix.

1. INTRODUCTION

JPEG is a widely used lossy compression standard (JPEG Standard, 1992) based on the Discrete Cosine Transform (DCT). It is appropriate for color images, as well as for grayscale ones. The image is first divided into blocks of size 8 x 8 pixels. In the following we will only encounter grayscale images, with pixel values ranging from 0, for pure black, to 255, for pure white. The DCT block computes a frequency map with 8 x 8 components, the first one representing the average value in each block and the others, the successively higher frequency within the block. The 64 DCT coefficients $c_{ij}, i = \overline{0,7}, j = \overline{0,7}$, correlate to corresponding frequencies, that is c_{00} correlates to the low frequencies of the original image block and the following DCT coefficients correlate to higher and higher frequencies of the image block. The elements of the matrix which results after DCT depend on the horizontal, diagonal and vertical frequencies, because of the cosine function used. Taking into account that the human eye is not sensitive to high frequencies, these can be discarded after the DCT is performed without affecting low frequency information. This is done in the quantization stage, when each of the 64 frequency components is divided by a separate quantization coefficient $q_{ij}, i = \overline{0,7}, j = \overline{0,7}$, and the results are rounded to the nearest integers.

The larger the quantization coefficients are, the more data is discarded. Higher frequencies are usually less accurately quantized than the lower ones, being less visible to the human eye. Also, the luminance data is usually more accurately quantized than chrominance data (Wallace, 1992), by using separate quantization tables. Usually, in this stage, the coefficients corresponding to higher frequencies result in 0, so that they are discarded, leading to the lossy part of compression

After the quantization stage, the reduced coefficients are encoded using either a Huffman or an arithmetic code. Then, the image is reconstructed through decompression, using the Inverse Cosine Transform (IDCT). The de-compressor multiplies the reduced coefficients by the elements in the quantization table

This paper was recommended for publication by Dorel Aiordachioaie

(entries), generating the reconstructed DCT coefficients. Although a great amount of DCT coefficients were discarded prior to block decompression, this does not lead to visible errors.

2. THE PROPOSED QUANTIZATION MATRIX MODIFICATION

In the quantization stage we can get varying levels of image compression and quality by selecting specific quantization matrices which have to be transmitted to the receiver. Subjective experiments lead to the JPEG standard quantization matrix, which accomplishes both high compression and good reconstructed image quality. This matrix corresponds to a quality level of 50 on a scale ranging from 1 to 100, where 1 gives the poorest image quality and highest compression, while 100 gives the opposite case. For a quality level greater than 50, the standard quantization matrix is multiplied by $\frac{100-quality\ level}{50}$, and for a quality level less than 50, the standard quantization matrix is multiplied by $\frac{50}{quality\ level}$ (Jain and Panchanathan, 1994; www.andrew.cmu.edu/user/slucey/dissert.pdf). Then the coefficients in the quantization matrix are rounded to the nearest integer. As we already have specified, the coefficients situated near the upper left corner of the image block correspond to lower frequencies to which the human eye is most sensitive. The quantized DCT matrix with elements $round\left(\frac{c_{ij}}{q_{ij}}\right)$ contains many zeros, which represent the less important higher frequencies that have been discarded. Only remaining nonzero coefficients will be transmitted and used to reconstruct the image.

In this method, for all quality levels, all the quantization steps are modified by the same factor, irrespective of the frequency content of the image.

A method to design the JPEG quantization matrix using rate-distortion approach and human visual system model is presented by Fong et al. (1997). Similarly, Berman et al. (1993) highlighted the effects of quantization matrix modification on cervical radiographs.

We will propose a way to modify differently the quantization step, according to the frequency content of the image. By taking the quantization coefficients q_{ij} in the standard quantization matrix from left to right and top to bottom, as shown in Fig. 1, we form an array, with elements $q_k, k=\overline{0,63}$.

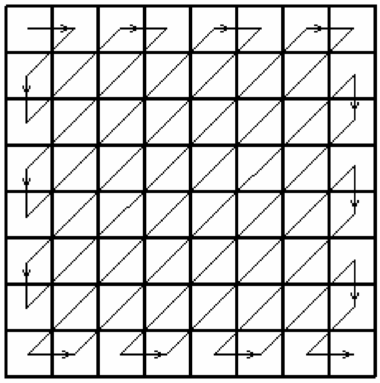

Fig. 1. Zig-zag scanning of 8 x 8 pixel image

For low frequency content images, the first L elements of this array are modified according to the relation

$$(1)\ qm_k = q_k \cdot n_1,\ 0 \le k \le L-1$$

and

$$(2)\ qm_k = q_k \cdot n_2,\ L \le k \le 63,$$

respectively, where n_1 is a quantity less than 1, chosen to decrease the quantization step for low frequency components and n_2 is a quantity greater than 1 chosen to increase the quantization step for high frequencies. L is a threshold value for the frequency range chosen so that, together with n_1 and n_2, leads to an improved compression for comparable values of error metrics and subjective human perception.

For high frequency content images, n_1 and n_2 act in opposite way as above, being quantities greater and less than 1, respectively. In this method we do not have to send the quantization matrix to the receiver, all we have to do is to send the correction parameters n_1, n_2 and the threshold *L*, the standard quantization matrix being known by the receiver. The threshold *L* is decided according to frequency map of the image.

Denoting by $x(i,j)$ the original image and by $y(i,j)$, its reconstructed version, the error metrics used to estimate the quality or fidelity of an image are (Sayood, 2001; Salomon, 2000):

- the Mean Square Error (MSE), given by

$$(3)\ MSE = \frac{1}{8\times 8}\sum_{i=1}^{7}\sum_{j=1}^{7}\left(x(i,j)-y(i,j)\right)^2$$

This measure gives an average value of the energy loosed in the compression of the original image.

- the Signal to Noise Ratio (SNR)

$$(4)\ SNR = 10 \cdot \log_{10} \frac{\frac{1}{8\times 8}\sum_{i=1}^{7}\sum_{j=1}^{7}\left(x(i,j)\right)^2}{\frac{1}{8\times 8}\sum_{i=1}^{7}\sum_{j=1}^{7}\left(x(i,j)-y(i,j)\right)^2}$$

SNR is measured in dBs and it gives indication about the signal level compared to the noise.

- the Peak Signal to Noise Ratio (PSNR)

$$(5)\ PSNR = 10 \cdot \log_{10} \frac{255^2}{\frac{1}{8\times 8}\sum_{i=1}^{7}\sum_{j=1}^{7}\left(x(i,j)-y(i,j)\right)^2}$$

PSNR is a more subjective qualitative measurement of distortion and finds the maximum signal to noise ratio (for a 8 bit image the maximum signal value is 255). PSNR is superior to other measures as it uses a constant value of the signal to compare to the noise, instead of a fluctuating one, as in SNR.

- the rate (bits per pixel)

$$(6)\ Rate(\text{bpp}) = \frac{\text{8 x Output File size}}{\text{Input File size}}$$

Individually, MSE, SNR and PSNR are not very good to indicate the subjective image quality, but used together, these error metrics are at least adequate to determine if an image is reproduced at a certain quality. It has been found (www.andrew.cmu.edu/user/slucey/dissert.pdf) from a subjective point of view that if a continuous tone image has a SNR value approximately above 30 dB, PSNR above 35 dB and MSE value below 20, the difference between the original and the reconstructed image is negligible to the viewer.

3. CASE STUDY

In our study we encounter two types of images, one having low frequency content and the other, high frequency content.

A. For the test image Zelda.bmp, with low frequency content, first we use the standard quantization matrix

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

and then the modified one, QM1, obtained for parameters L=8, $n_1 = 0.8$, $n_2 = 2.5$.

$$QM1 = \begin{bmatrix} 13 & 9 & 8 & 13 & 60 & 100 & 128 & 153 \\ 10 & 10 & 12 & 48 & 66 & 145 & 159 & 138 \\ 12 & 33 & 40 & 60 & 100 & 143 & 173 & 140 \\ 35 & 43 & 55 & 73 & 128 & 218 & 200 & 155 \\ 45 & 55 & 93 & 140 & 170 & 255 & 255 & 193 \\ 60 & 88 & 138 & 160 & 203 & 255 & 255 & 230 \\ 123 & 160 & 195 & 218 & 255 & 255 & 255 & 253 \\ 180 & 230 & 238 & 245 & 255 & 250 & 255 & 248 \end{bmatrix}$$

The standard quantization coefficients are plotted in Fig. 2 with dashed line and the modified coefficients with solid line.

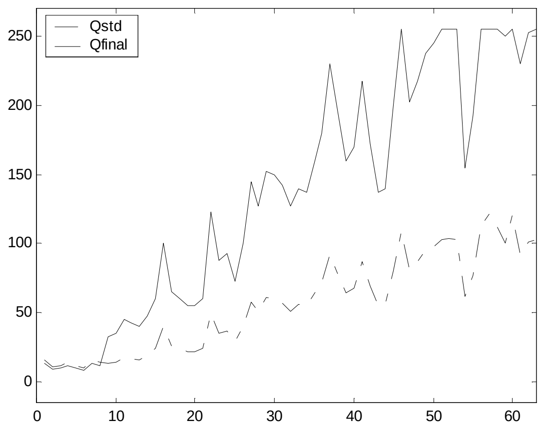

Fig. 2. Coefficients of the standard and the modified quantization matrix

The two images obtained by compression with matrices QM and $QM1$ are given in Figs. 3a and 3b, respectively.

Fig. 3a

Fig. 3b

The error metrics computed for the image Zelda.bmp compressed with the standard quantization matrix Q and the modified matrix QM1, respectively, are given in Table 1.

Table 1 Eror metrics for image Zelda.bmp compressed by means of standard quantization matrix Q and the modified quantization matrix QM1

Image	MSE	SNR	PSNR	Rate%	Rate (bpp)
Zelda Q	19.49	27.08	35.23	4.78	0.38
Zelda QM1	20.05	26.98	34.93	4.52	0.36

From data in Table 1, we see that the loss in MSE for the images compressed with QM1 a is 2%, the loss in SNR is 1 % and the loss in PSNR are 0.8%, respectively, against the standard matrix. This causes no subjective loss in image quality, while the gain in bits per pixel is 5.5%

B. Another image we analyze is the test image Baboon.bmp that is a high contrast image, with a lot of high frequencies, with a non-uniform area distribution, because the stripes near the nose. The parameters we chose are L=18, $n_1 = 1.4$, $n_2 = 0.71$. The quantization matrix $QM2$ obtained for these parameters is:

$$QM2 = \begin{bmatrix} 22 & 15 & 14 & 22 & 34 & 56 & 36 & 44 \\ 17 & 17 & 20 & 27 & 36 & 41 & 43 & 39 \\ 20 & 18 & 22 & 34 & 29 & 41 & 49 & 40 \\ 20 & 24 & 16 & 21 & 36 & 62 & 57 & 44 \\ 25 & 16 & 26 & 40 & 49 & 78 & 74 & 55 \\ 17 & 25 & 39 & 46 & 58 & 74 & 81 & 66 \\ 35 & 46 & 56 & 62 & 74 & 86 & 86 & 72 \\ 51 & 66 & 68 & 70 & 80 & 71 & 74 & 71 \end{bmatrix}$$

The standard quantization coefficients are plotted in Fig. 4 with dashed line and the modified coefficients with solid line.

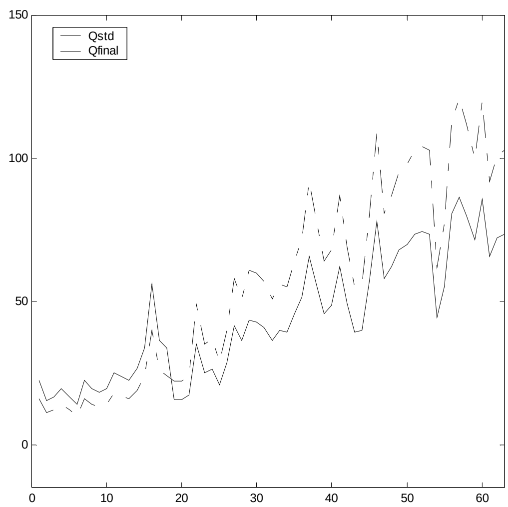

Fig. 4. Coefficients of the standard and the modified quantization matrix

The two images obtained by compression with matrices Q and $QM2$ are given in Figure 5a and b, respectively.

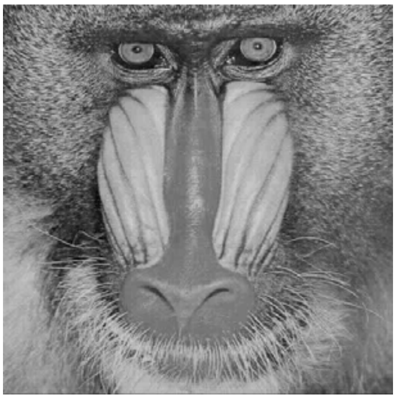

Fig. 5a

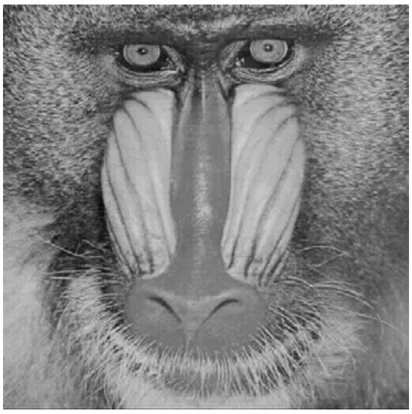

Fig. 5b

The quantitative evaluation of error metrics computed for images in Figs. 5a and b are given in Table 2.

Table 2. Error metrics for image Baboon.bmp compressed by means of standard quantization matrix Q and the modified quantization matrix QM2

Image	MSE	SNR	PSNR	Rate %	Rate (bpp)
Baboon Q	156.8	20.70	26.17	12.89	1.031
Baboon QM2	155.7	20.73	26.20	12.73	1.013

From the data in Table 2, we see that all error metrics are improved (0.7% for MSE, 0.1 % for SNR and PSNR), and the gain in bits per pixel comparing to the case when the standard matrix is used is 5%.

4. CONCLUSIONS

In our work we propose a way to refine the standard quantization matrix used in JPEG method according to the frequency content of the image, by applying a correction factor, so that for images with low frequency content, the quantization step decreases for low frequencies and increases for higher ones and for images with high frequency content the correction acts in opposite way.

5. REFERENCES

JPEG Standard (1992) *Information technology – digital compression and coding of continuous-tone still images – requirements and guidelines*, Recommendation T.81, CCITT T.81.

Wallace, G.K. (1992). The JPEG still picture compression standard. IEEE Trans. Consumer Electronics, **38**, No.1, Feb. pp. xviii– xxxiv.

Jain, A., and S. Panchanathan, (1994). Scalable compression for image browsing. IEEE Trans. Consumer Electronics, **40**, Aug pp. 394 – 404.

http://www.andrew.cmu.edu/user/slucey/dissert.pdf

Fong, W. C., S. C. Chan, and K. L. Ho, (1997). Designing JPEG quantization matrix using rate-distortion approach and human visual system model, *ICC 97 Montreal IEEE International Conference on Communications*, *1997*. Vol. **3**, pp. 1659 – 1663.

Berman, L. E., R. Long, and S. R. Pillemer, (1993). Effects of Quantization Table Manipulation on JPEG Compression of Cervical Radiographs, *International Symposium, Seminar, & Exhibition, Seattle,* WA, May 16-21.

Sayood, K. (2001) *Introduction to data compression*, Morgan Kaufman.

Salomon, D. (2000). *Data compression. The complete reference.* Springer Verlag.