CREATIVE MATH.
12 (2003), 00-00

The limitations of JPEG compression

Ovidiu Cosma

Abstract. The JPEG (*Joint Photographic Experts Group*) compression standard is based on the Discrete Cosine Transform. This transform was chosen for efficiency, but has the major disadvantage that it does not localize the frequency components in space. Because of that, the image has to be partitioned prior transform, and the margins of the blocks become visible at high compression ratios. This problem is solved in the JPEG 2000 standard, which uses the Discrete Wavelet Transform.

1. The JPEG standard

The JPEG standard was developed for compressing static images with continuous tones and has the following four operation modes [9], [2]:

1. Sequential, in which the image is coded in a single pass, in the left - right - down order.
2. Progressive, in which the image is coded in several passes. This type of coding is useful in the case that the images are transmitted through computer networks, and the receiver wants to watch the image reconstructing process while receiving data.
3. Lossless, that has the advantage that the original image can be exactly reconstructed, but with the cost of a much lower compression ratio.
4. Hierarchical, in which the image is coded at multiple resolutions, and the low-resolution versions can be decoded independently of the high resolution ones.

1.1. **The sequential coding mode.** This operation mode is illustrated in figure 1, and contains more transformation stages. The image preparation module performs the conversion of the image from the RGB colour space to a *luminance – chrominance* (LC) colour space [1]. The LC colour spaces are more suitable for image compression, because they enable the speculation of the human eye characteristics, that is much more sensitive to light than colour. In consequence the chrominance components can be represented at lower resolutions than the luminance component, without inducing perceptible distortions. The chrominance components resolution is halved in both the horizontal and vertical directions. Thus the total amount of data is halved, without affecting the quality of the images.

Received: 17.05.2003; In revised form:
Key words and phrases. JPEG image compression.

Next the image is partitioned uniformly in 8 x 8 pixel blocks, and the *Discrete Cosine Transform* (DCT) [10] is applied for each of the blocks. Relation (1) indicates the DCT computing process for a pixel block $f(x, y)$. The inverse transform is given in (2).

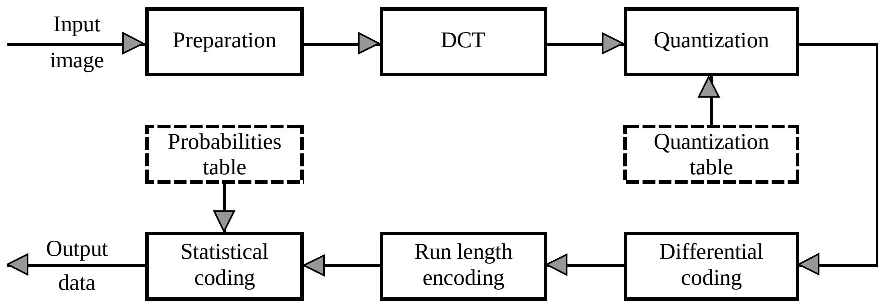

Figure 1: The block scheme of a JPEG coder, in the sequential mode of operation

$$F(m,n) = \frac{1}{4}\sqrt{\lambda_{mn}}\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)\cos[\frac{m\pi}{8}(x+\frac{1}{2})]\cos[\frac{n\pi}{8}(y+\frac{1}{2})] \qquad (1)$$

$$f(x,y) = \frac{1}{4}\sum_{m=0}^{7}\sum_{n=0}^{7} \sqrt{\lambda_{mn}}F(m,n)\cos[\frac{m\pi}{8}(x+\frac{1}{2})]\cos[\frac{n\pi}{8}(y+\frac{1}{2})] \qquad (2)$$

$$\lambda_{mn} = \begin{cases} 1/4, & \text{if} \quad m = n = 0 \\ 1/2, & \text{if} \quad m > 0, n = 0 \text{ or } m = 0, n > 0 \\ 1, & \text{if} \quad m > 0, n > 0 \end{cases}$$

After this transformation, the overall average will be placed in the upper-left corner, and the high frequency components coefficients will be placed in the lower-right part of the blocks. The transformation stage induces an approximation error because of the limited number of coefficients, and because of the inexact representation of their value.

This error increases in the quantization process that has the task to eliminate the less important DCT coefficients. A quantization table that defines a threshold for each position in the 8 x 8 coefficient matrix is used for this purpose. The coefficients whose absolute value is under the corresponding threshold will be eliminated in the process.

The quantization table is not standardized, allowing the competition between different implementations. The thresholds in this table establish the compression ratio and the distortion level. The quantization table must take into account the importance of each of the frequency components of images. Because the high frequency components of images can be hardly distinguished by the human eye, they can be roughly approximated. Those components usually have low amplitudes in the case of natural images. In consequence there is a good probability that they will be eliminated in the quantization process and zeros will accumulate in the lower-right part of the blocks. There have been proposed several types of quantization tables, which have been optimized experimentally.

In the next stage, the differential coding [2] of the block averages is performed. Thus the first coefficient of every block is subtracted from the first coefficient of the previous block and only the difference is retained. Usually the block averages have relatively large values, but they alter slightly from one block to another. Because of this property is expected that after the differential coding will result small differences or even zeroes.

No compression was achieved until now, but the entropy of the image [7] was reduced in the DCT and quantization blocks. The previous operations had the task to prepare the ground for the next module that performs the *run length encoding* [2].

In order to benefit from the zeroes accumulated in the quantization process, the DCT coefficients will be placed in the output stream by crossing each of the blocks in zigzag order. There is a high probability that because of this rule long runs of zeroes will appear. This process is illustrated in figure 2.

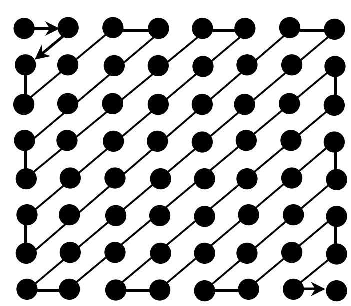

Figure 2: Placing the DCT coefficients in the output stream, by crossing the blocks in zigzag order

The last of the modules performs a statistical coding of data [2]. The statistical coding requires that the application specify one ore more probability tables. They can be predefined or computed in a separate stage prior compression, using a set of representative images. If there are more probability tables, an index has to be included in the output stream, for specifying the one that was used in the coding process. If the table is computed by the coding application, it must be inserted in the output stream.

If an adaptive coding is used (Huffman [6] or arithmetic [3]) there is no need for external tables, because the image statistics can be updated during the coding process. The arithmetic coding is less efficient, but increases the compression scheme performance with up-to 5 – 10%.

2. Performance evaluation

The most important disadvantage of the image compression schemes that are based on the DCT is the fact that the transform does not reveal any information about the space localization of the frequency components. Because of that, images must be partitioned in blocks that are transformed separately. At high compression ratios, the approximations performed in the quantization step can create important differences between the neighbouring pixels, at the edge of the blocks. Those differences correspond to horizontal and vertical steps in the luminance and chrominance components. This type of distortion is highly visible especially in the long smooth areas with continuous tones, and because of it the images loose their natural aspect.

The bit rate at which this effect appears depends on the characteristics of the image, but if it is decreased any further, the visual quality of images worsens very quickly.

A much better transform for image compression is the *Discrete Wavelet Transform*, (DWT) because it localizes the frequency components in space and does not require the partitioning of images. This type of transform is used in the JPEG 2000 compression standard [5].

In figures 3 and 4 a comparison between the performance of a JPEG coder and a DWT based image compression scheme is presented. The graphs were generated for a usual image, with reasonable amounts of details and smooth areas. The DWT compression scheme uses the HIA [1] and SPIHT [8] encodings in order to increase performance. The Villasenor 18/10 filter [4] was used for computing the DWT.

In figure 3, "A" marks the point where the margins of the blocks became visible in the JPEG compression graph, and "B" marks the point where the reconstructed image totally looses it's natural aspect. In the case of the DWT compression scheme, the reconstructed image has good visual quality even at 0.2bpp. The PSNR is situated at this bit rate somewhere between points "A" and "B" in the JPEG graph, but the distortions are not near so annoying.

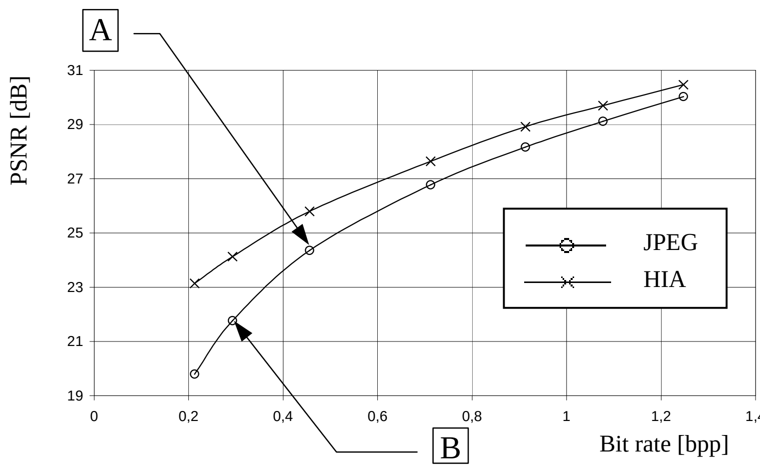

Figure 3: The evolution of the PSNR (*Peak Signal-to-Noise Ratio*) as a function of bit rate, in the case of JPEG and Wavelet SPIHT - HIA compression

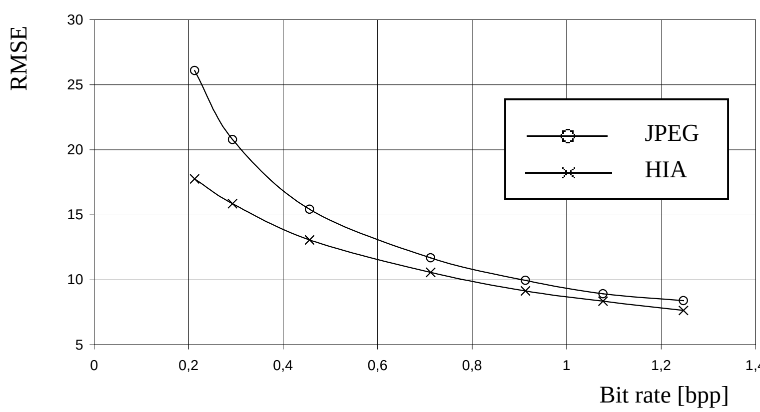

Figure 4: The evolution of the RMSE (*Root Mean Square Error*) as a function of bit rate, in the case of JPEG and Wavelet SPIHT - HIA compression

References

[1] Cosma O., *Contributions to the Coding of Image Subbands*, thesis, Politehnica University Bucharest, 2003.

[2] Cosma O., *Multimedia*, course, North University of Baia Mare, 2000.

[3] Howard P.G., Vitter J.S., *Arithmetic Coding for Data Compression*, Proceedings of the IEEE, **82**(6), 06.1994.

[4] http://www.isye.gatech.edu/∼brani/.public_html/Wiley/filters.html, *A Bank of Wavelet Filters.*

[5] ISO/IEC, *JPEG 2000 part I, final committee draft version 1.0*, ISO/IEC JTC 1/SC 29/WG 1, 03.2000.

[6] Low J., *Adaptive Huffman Coding*, www.cs.duke.edu/csed/curious/ compression/adaptivehuff.html

[7] Marshal D., *Lossless Compression Algorithms (Entropy Encoding)*, http://www.cs.cf.ac.uk/Dave/Multimedia/node207.html

[8] Said A., Pearlman W.A., *A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees*, IEEE Transactions on Circuits and Systems for Video Technology, **6**, june 1996.

[9] Wallace G.K., *The JPEG Still Picture Compression Standard*, Communications of the ACM, **34**, Issue 4, 04.1991.

[10] Wikipedia, *Discrete Cosine Transform*, http://www.wikipedia.org/wiki/Discrete_ cosine_transform

Ovidiu Cosma
Department of Mathematics and Computer Science
Faculty of Sciences
North University of Baia Mare
Victoriei 76, 430122 Baia Mare, Romania
E-mail address: cosma@alphanet.ro