CREATIVE MATH. & INF.
17 (2008), No. 2, 107 - 114

Online version available at http://creative-mathematics.ubm.ro/
Print Edition: ISSN 1584 - 286X Online Edition: ISSN 1843 - 441X

Image compression with a human touch

OVIDIU COSMA

ABSTRACT.
Compression is very important for multimedia applications, because it reduces the amount of space for depositing the information, and the bandwidth required for sending it through a computer network. Multiresolution analysis performs a wavelet decomposition that decorrelates the signal data, in preparation for the quantization step and the final redundancy reduction. This article presents a simple algorithm for performing the decomposition of images and propounds a quantization method that takes into account the human eye contrast sensitivity function.

1. INTRODUCTION

An image compressor is usually composed of an entropy reduction block, followed by a redundancy reduction algorithm. The *Entropy Reduction* block applies a series of transformations to the signal, that are not fully reversible, in order to prepare the data for the stage of *Redundancy Reduction*. The *Entropy Reduction* block contains two main components: *Transformation* and *Quantization*, as it is shown in Figure 1.

$$\overbrace{\boxed{\text{Transformation}} \to \boxed{\text{Quantization}}}^{\text{Entropy reduction}} \to \boxed{\text{Redundancy reduction}}$$

Figure 1

The transformation stage has the task of conveying the image in a domain where its important traits can be easily marked out. Most of the actual compressors use the *Fourier Transform* (FT) for this purpose. The FT is a poor choice for image compression, because it only reveals information about the spectral components of the signal, and it gives no information about the time localization of these components. Because of that, dissimilar signals can have similar transforms, and if the transform is followed by quantization, the little components that make the difference can disappear [1].

We can say that the FT has the best possible frequency resolution, and the worst time resolution, and it is useful in the case of unstationary signals only for revealing their spectral components [1]. If time localization of the spectral components is needed, the FT is not the best choice. However, the FT is used in the old *JPEG* and *MPEG* compression standards [4]. The new *JPEG 2000* standard uses the wavelet transform [5].

The *Quantization* step is responsible for eliminating the less significant information from the transformed image. The compromises performed at this stage determine the distortion level of the reconstructed image, and the compression ratio [4].

Redundancy Reduction is the final stage that benefits from the previous transformations, and actually achieves compression. It usually performs a *run length encoding* followed by *Huffman* or *Arithmetic Encoding* [4]. Other compression techniques are described in [2] and [3].

2. IMAGE COMPRESSION

The purpose of the transform stage is to obtain a reversible decomposition of the signal in a set of basis components. The FT uses sinusoides and the WT uses shorter components. Let's consider the job done, and the signal $f(t)$ expressed as a weighted sum of basis functions: $\psi_1(t), ..., \psi_m(t)$.

$$f(t) = \sum_{i=1}^{m} c_i \psi_i(t) \tag{2.1}$$

The data set required for the reconstruction of the signal contains the coefficients $c_1, ..., c_m$. In order to achieve compression, an approximation of $f(t)$ must be found, that can be expressed with fewer coefficients. For that purpose, the coefficients $c_1, ..., c_m$ are sorted in order of significance, so that for every $\tilde{m} < m$, the first $\tilde{m}$ elements of the sequence give the best approximation $\tilde{f}(t)$ of $f(t)$ in the L^2 norm.

The solution to this problem is simple if an orthonormal basis is used. Let σ be a permutation of the elements $1, ..., m$, and $\tilde{f}(t)$ an approximation that uses coefficients corresponding to the first $\tilde{m}$ numbers of this permutation:

Received: 11.07.2008; In revised form: 12.11.2008.; Accepted:
2000 *Mathematics Subject Classification*. 94A08.
Key words and phrases. *Image compression, numeric codes.*

$$\widetilde{f}(t)\sum_{i=1}^{\widetilde{m}} c_{\sigma(i)}\psi_{\sigma(i)}(t) \tag{2.2}$$

The square of the approximation error in the L^2 norm is:

$$\begin{aligned}
&\left\|f(t)-\tilde{f}(t)\right\|_2^2 = \left\langle f(t)-\tilde{f}(t) | f(t)-\tilde{f}(t)\right\rangle = \\
&= \left\langle \sum_{i=\tilde{m}+1}^{m} c_{\sigma(i)}\psi_{\sigma(i)} \;\middle|\; \sum_{j=\tilde{m}+1}^{m} c_{\sigma(j)}\psi_{\sigma(j)} \right\rangle = \\
&= \sum_{i=\tilde{m}+1}^{m}\sum_{j=\tilde{m}+1}^{m} c_{\sigma(i)}c_{\sigma(j)}\left\langle \psi_{\sigma(i)}|\psi_{\sigma(j)}\right\rangle = \\
&= \sum_{i=\tilde{m}+1}^{m} (c_{\sigma(i)})^2
\end{aligned} \tag{2.3}$$

The last step is possible only if the basis is orthonormal, that means that

$$\langle\psi_u|\psi_v\rangle = \delta_{uv}$$

where δ_{uv} is the *delta Kronecker* function, that has the value 1 for $u = v$, and 0 in rest.

In conclusion, for minimizing the approximation error, for every $\tilde{m}$, the best choice for σ is the permutation that sorts the coefficients in descending order [11], that is

$$\left|c_{\sigma(1)}\right| \geqslant \left|c_{\sigma(2)}\right| \geqslant \cdots \geqslant \left|c_{\sigma(m)}\right| \tag{2.4}$$

3. The Haar transform

The Haar transform performs the decomposition of a signal in a set of Haar wavelets. The Haar wavelets are defined in [5], and they are obviously the simplest wavelet functions.

$$\psi_{jk}(x) = \psi(2^j x - k), \quad k = 0, \ldots, 2^j - 1$$

where

$$\psi(x) = \begin{cases} 1, & \text{for } 0 \leqslant x < \dfrac{1}{2} \\ -1, & \text{for } \dfrac{1}{2} \leqslant x < 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

The next figure illustrates the form of the Haar wavelets.

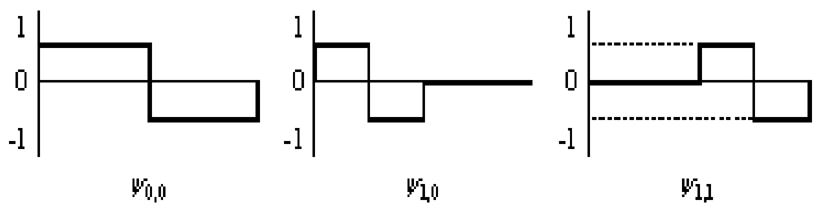


Figure 2

It is obvious that the Haar wavelets have compact support and they are orthogonal, but for image compression it is important that they are also normalized. The following relation defines the normalized Haar wavelets:

$$\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k) \tag{3.2}$$

The wavelets defined by relation (3.2) are orthonormal. This implies that they have the property expressed in (2.3) [7].

In order to compute the Haar transform of a digital signal $f(i)$, we start with applying the following filter, that splits the signal spectrum in two equal parts:

$$h_1(i) = (f(2i) - f(2i+1))/\sqrt{2}$$
$$l_1(i) = (f(2i) + f(2i+1))/\sqrt{2} \tag{3.3}$$

Where $h_j(i)$ contains the high frequencies coefficients, and $l_j(i)$ contains the averages at the current resolution. Because natural and computer generated images are usually composed by long smooth portions separated by localized discontinuities, we are satisfied with the sharp time resolution of the $h_1(i)$ signal, and we continue with the decomposition in the same manner of the average signal $l_j(i)$, $j=1,\ldots,n$. The process is completed when the signal is reduced to a single sample, that will contain the overall average of the original signal.

This process is illustrated in the following figure:

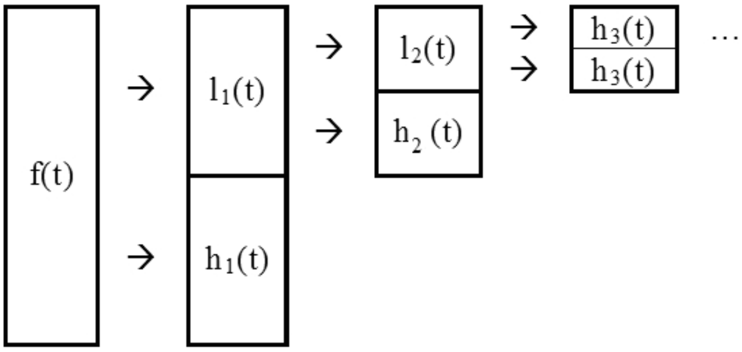


Figure 3

For computing the inverse transform, we start with the general average and the detail coefficient for the lowest resolution, and we determine the two averages for the next (higher) resolution.

$$l_j(2i) = (l_{j+1}(i) + h_{j+1}(i))/\sqrt{2}$$
$$l_j(2i+1) = (l_{j+1}(i) - h_{j+1}(i))/\sqrt{2}, \quad j = n-1, ..., 0 \tag{3.4}$$

The transform is repeated until the original signal $f(t)$ is recovered. $f(i) = l_0(i)$.

4. Two-dimensional Haar Wavelet transform

There are two possibilities to extend a one-dimensional wavelet transform to two-dimensional signals:

- Perform a complete separate one-dimensional wavelet transform for each of the rows, and then the collection of transformed rows is interpreted as a new image, and compute the wavelet transform for each of the columns [1],[6].
- Alternate between the operations on the rows, and the operations on columns. At each step only the averages resulted from the previous operations are processed.

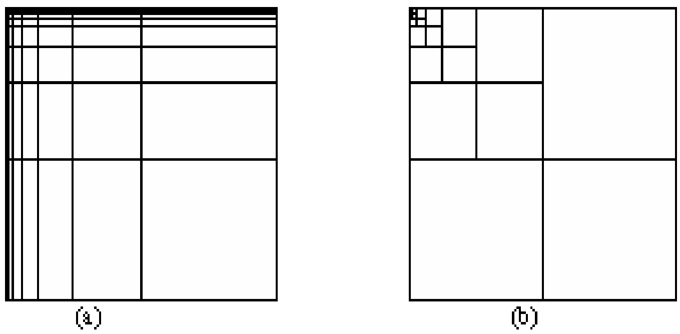


Figure 4

The second method is more efficient, because it involves fewer operations.
Figure 4 presents the localization of the subbands generated by these decomposition methods (first method in a and second in b). The right – down square contains the highest detail coefficients, and the upper – right corner contains the overall average.

5. Quantization

The role of the quantization step is to eliminate the insignificant information from the transformed image. A general method of quantization that can be applied in the case of wavelet transform was described in [4]. There are many ways to perform the quantization process. The most common method is the use of a linear quantizer with a given quantization step ΔQ, so that the quantized representation $\tilde{c}$ of a coefficient c is given by:

$$\widetilde{c} = \Delta Q \left\lfloor \frac{c}{\Delta Q} \right\rfloor \tag{5.1}$$

The linear quantization method is correct from a mathematical point of view, but it does not take in consideration the human eye characteristics.

The human eye contrast sensitivity function CSF is presented in the next figure. The luminance CSF has a maximum for spatial frequencies of around 4 cycles per optical degree (cpd). It is important to know which subband contains this maxim. For a viewing distance of between 30 and 90 cm the maxim is situated in levels 3-5. A weighting computed for a small distance would coarsely quantizate the coefficients in level 5, but for larger distances this components shift to the maximum sensitivity region [8]

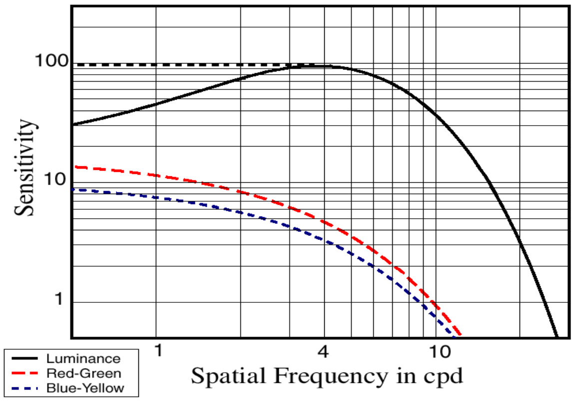


Figure 5

To avoid this situation, the luminance CSF is flattened for low frequencies as shown with the horizontal dotted line in Figure 5.

Because the wavelet decomposition performs a natural separation of the coefficients per subbands, it is a good idea to apply a separate weighting factor for each subband, before the quantization step. The weighting factors are chosen in concordance with the CSF.

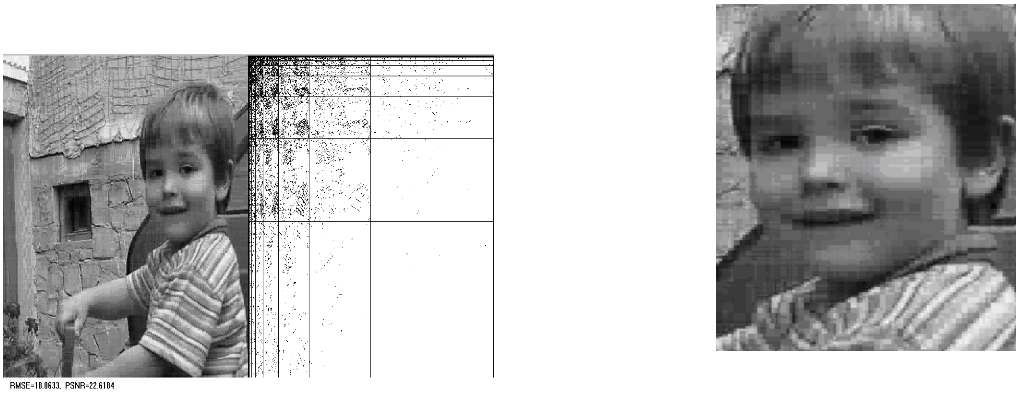
Figure 8 presents two sample images generated with the algorithms described in this article, and close-ups that highlight the artifacts generated by the transforms. For each of the pictures, the localization of the significant wavelet coefficients is shown. The black dots indicate the position of the nonzero coefficients, and the white dots indicate the null ones. The distortion level is the same for both images: PSNR $\approx$ 22,61, and RMSE $\approx$ 18,86. RMSE is the Root Mean Square Error and PSNR is the Peak Signal – to – Noise Ratio (in dB):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - \tilde{f}_i)} \qquad \text{PSNR} = 20\log_{10}\frac{\max_i |f_i|}{\text{PSNR}} \tag{5.2}$$

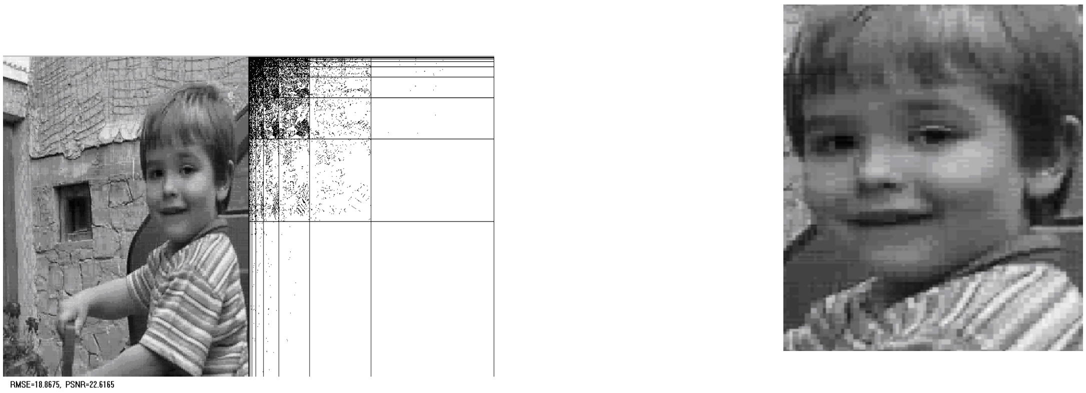
where N is the total number of pixels, and $\max_i |f_i|$ is the difference between the maximum and the minimum value the pixels can hold.

For the image in a, a linear quantization was used, and for the image in b the weighting factors presented in Table 1 were applied, in concordance with the CSF. Although the second image has fewer coefficients, and the distortion levels are equal, it seems to have a better quality than the image in a, due to the CSF.

Level	Weighting
1	0,16
2	0,29
3	0,56
4	0,89
5	1
6	1
...	



(a)



(b)

Figure 6

References

[1] Cosma, O., *Wavelet Transform application in image compression*, Buletinul Ştiinţific al Universităţii din Baia Mare seria B, 2000
[2] Cosma, O., *O metodă de compresie a imaginilor,* Prima conferinţă de matematici aplicate, Baia Mare, 1998
[3] Cosma, O., *Extensions for the Image Broadcasting Protocol*, Buletinul Ştiinţific al Universităţii din Baia Mare seria B, 1997
[4] Cosma, O., *Multimedia,* Univ. de Nord Baia Mare, 2000
[5] ISO/IEC, JPEG 2000 part I, final committee draft version 1.0
[6] Fournier, A., *Wavelets and their Applications in Computer Graphics*, SIGGRAPH Course Notes, University of British Columbia 1995
[7] Walker, S. J., *Fourier Analysis and Wavelet Analysis,* Notice of the AMS, Vol. 44, No. 6
[8] Nadenau, M., *Integration of Human Color Vision Models into High Quality Image Compression,* Ecole Polytechnique Federale de Lausane, 2000
[9] Hilton, L.M., *Compressing Still and Moving Images With Wavelets,* Multimedia Systems, Vol. 2, No. 3, 1994
[10] Stollniz, J. E., *Wavelets for Computer Graphics*, IEEE Computer Garphics and Applications 05. 1995
[11] Graps, A., *An Introduction to Wavelets*, IEEE Computational Science and Engineering, Vol. 2, No. 2, 1995

Noth University Of Baia Mare
Department of Mathematics and
Computer Science
Victoriei 76
430120 Baia Mare, Romania
E-mail address: cosma@mail.alphanet.ro