Mathematical and Computer Modelling 52 (2010) 1265–1270

Contents lists available at ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

Wavelet Transform application to the compression of images

Macarena Boix [a], Begoña Cantó [b,*]

[a] Departamento de Matemática Aplicada, Universidad Politécnica de Valencia, Escuela Politécnica Superior de Alcoy, Plaza Ferrándiz y Carbonell 2, 03801 Alcoy (Alicante), Spain

[b] Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia, 46022 Valencia, Spain

ARTICLE INFO

Article history:
Received 29 September 2009
Accepted 5 February 2010

Keywords:
Wavelet Transform
Compression images
Global threshold

ABSTRACT

In this paper a method for image compression is described. In the Wavelet Transform technique the coefficients below a certain threshold are removed. We propose a new type of global threshold to improve the wavelet compression technique. The aim is to maintain the retained energy and to increase the compression ratio with respect to other global thresholds commonly used.

1. Introduction and background

The wavelet compression technique applied to images may be often useful in telemedicine to reduce the bandwidth needed for transmitting graphics files for analysis or remote diagnostics.

Images are formed by pixels that are little pictures whose color is associated to a numerical value according to the classification used. The information obtained is stored in one or more matrices (Fig. 1).

Computer images are extremely data intensive and hence require large amounts of memory for storage. The compression technique reduces the size of a file in order to facilitate the efficient transfer of their storage. Basically we can distinguish between a lossless compression with rates of 3:1 and a lossy compression with much higher rates, even 30:1.

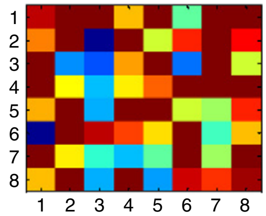
Lossless compression is used when data before and after compression are accurate. This technique is used when binary data such as executables, documents etc. are compressed. On the other hand, images, music, video and so on cannot be reproduced exactly. In the images case, an approximation of the original image is enough. In this case, lossy compression techniques must be used for the sake of using a minimum storage space. They consist of removing the redundant information in adjacent pixels to minimize the number of bits.

Wavelet techniques have been proven to be indispensable for image processing. For that, the Wavelet Transform is used in this work as the compression technique. Cite for example in 1992 the FBI used wavelet compression to store and retrieve its 250 million fingerprint files requiring 250 terabytes of space. Using wavelets, the FBI obtains a compression ratio of about 20:1. Note that the storage of a fingerprint with a gray scale of 8 tones and with a resolution of 500 dots per inch occupies 10 megabytes [1].

The Discrete Wavelet Transform (*DWT*) uses Mallat's multiresolution analysis to represent a signal in terms of its frequency components. *DWT* projects a signal on different approximation spaces, called $\mathbf{V}_j$ with different levels of resolution. Each of these includes another approach at half resolution. The difference between the two is captured by the spaces $\mathbf{W}_j$, called the subspaces of detail. If the signal behaves smoothly, the difference between two successive projections of the signal is very small and therefore the detail coefficients will be small or zero. These small differences are often imperceptible the eye (for more information see, for example, [2] and [3]).

* Corresponding author.
E-mail addresses: mboix@mat.upv.es (M. Boix), bcanto@imm.upv.es, bcanto@mat.upv.es (B. Cantó).

0895-7177/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2010.02.019



$$A = \begin{bmatrix} 60 & 92 & 81 & 44 & 68 & 30 \\ 48 & 73 & 1 & 93 & 37 & 54 \\ 89 & 17 & 13 & 46 & 83 & 15 \\ 76 & 40 & 20 & 41 & 50 & 69 \\ 45 & 93 & 19 & 84 & 70 & 37 \\ 2 & 91 & 60 & 52 & 42 & 86 \end{bmatrix}$$

Fig. 1. Pixels of different colors represented by a matrix.

Now, we give some mathematical tools used in this paper. Lebesgue spaces are given by:

$$\mathbb{L}^p(\mathbb{R}) = \left\{ x : \|x\|_p = \left(\int_{\mathbb{R}} |x(t)|^p \mathrm{d}t \right)^{1/p} < \infty \right\}, \quad p \geq 1$$

A Hilbert space $\mathbb{L}^2(\mathbb{R})$ is formed by the functions which have a finite norm. In these functions the energy $E(f)$ is defined as the square of the norm.

A sequence $\{V_j\}_{j\in\mathbb{Z}}$ of the closed subspaces of $\mathbb{L}^2(\mathbb{R})$ is a multiresolution approximation (MRA) if the following properties are satisfied:

(1) $\forall (j,n) \in \mathbb{Z}^2 \quad f(t) \in V_j \longleftrightarrow f(t-2^j n) \in V_j$.
(2) $\forall j \in \mathbb{Z} \quad V_{j+1} \subset V_j$.
(3) $\forall j \in \mathbb{Z} \quad f(t) \in V_j \longleftrightarrow f(t/2) \in V_{j+1}$.
(4) $\lim_{j\to-\infty} V_j = \overline{\bigcup_{j\in\mathbb{Z}} V_j} = \mathbb{L}^2(\mathbb{R})$.
(5) $\lim_{j\to+\infty} V_j = \bigcap_{j\in\mathbb{Z}} V_j = \{0\}$.
(6) There exists $\phi(t) \in V_0$ (scaling function) such that $\phi(t-n)_{n\in\mathbb{Z}}$ is a orthonormal basis of V_0.

For each integer j, the functions are an orthonormal basis for each V_j. From the MRA, we can express $\mathbb{L}^2(\mathbb{R}) = \cdots + V_{-1} + V_0 + V_1 + \cdots$, but this is not an orthogonal decomposition and then we introduce the wavelet spaces W_j from the orthogonal complements of V_j in V_{j-1} ($V_{j-1} = V_j \bigoplus W_j, \forall j \in \mathbb{Z}$). Moreover, we can construct wavelets ψ such that the dilated and translated family $\{\psi_{jn}(t) = 2^{-j/2}\psi(2^{-j}t - n)\}_{(j,n)\in\mathbb{Z}^2}$ is an orthogonal basis of $\mathbb{L}^2(\mathbb{R})$.

Taking into account that $\mathbb{L}^2(\mathbb{R}) = \overline{\bigoplus_{j=-\infty}^{\infty} W_j}$, for any arbitrary J, a finite energy signal $x(t)$ can be expressed in successive projections in the mentioned subspaces (approximations and details respectively):

$$x(t) = \sum_{n\in Z} a_{Jn}\phi_{Jn}(t) + \sum_{i=-\infty}^{J} \sum_{n\in Z} d_{in}\psi_{in}(t)$$

To construct wavelets in $\mathbb{R}^2$ we can use the technique of tensor product. Mallat's multiresolution analysis decomposes an image into a set of approximation coefficients (low frequency components) and the scale dependent hierarchy of the detail coefficients (high frequency components). A standard tensor product orthogonal wavelet transformation of an image results in three sets of generated detail coefficients: diagonal, horizontal and vertical.

The 2-D wavelet basis functions are constructed via the translations and dilations of a tensor product of univariate wavelets and scaling functions,

$$\begin{aligned} \phi(x_1,x_2) &= \phi(x_1)\phi(x_2) \\ \psi^h(x_1,x_2) &= \phi(x_1)\psi(x_2) \\ \psi^v(x_1,x_2) &= \psi(x_1)\phi(x_2) \\ \psi^d(x_1,x_2) &= \psi(x_1)\psi(x_2) \end{aligned}$$

where $\psi(\cdot)$ represents the shifted and dilated versions of a wavelet mother function and it is associated to a one-dimensional multiresolution approximation $\{\mathbf{V}_j\}_{j\in\mathbb{Z}}$ and $\phi(\cdot)$ represents the shifted versions of a low-pass scaling function. Let $h,\ v,\ d$ be the horizontal, vertical and diagonal directions, respectively and the expressions for $\phi_{j,\mathbf{k}}(x)$ and $\psi^i_{j,\mathbf{k}}(x)$, for $i = h, v, d$ and where $\mathbf{x} = (x_1,x_2) \in \mathbb{R}^2$, $\mathbf{k} = (k_1,k_2) \in \mathbb{Z}^2$ and are given by

$$\begin{aligned} \phi_{j,\mathbf{k}}(\mathbf{x}) &= 2^{2j}\phi(2^j x_1 - k_1, 2^j x_2 - k_2) \\ \psi^i_{j,\mathbf{k}}(\mathbf{x}) &= 2^{2j}\psi^i(2^j x_1 - k_1, 2^j x_2 - k_2). \end{aligned}$$

Let $\{\mathbf{V}^2_j\}_{j\in\mathbb{Z}}$ be defined by $\mathbf{V}^2_j = \mathbf{V}_j \otimes \mathbf{V}_j$ and let $\mathbf{W}^2_j$ be the detail space equal to the orthogonal complement of the lower resolution approximation space $\mathbf{V}^2_j$ in $\mathbf{V}^2_{j-1}$. To construct a 2-D wavelet basis, we give the following result (see [4] for more details).

Theorem 1. *Let ϕ be a scaling function and ψ be the corresponding wavelet generating a wavelet orthonormal basis of $\mathbb{L}^2(\mathbb{R})$. We define three wavelets:*

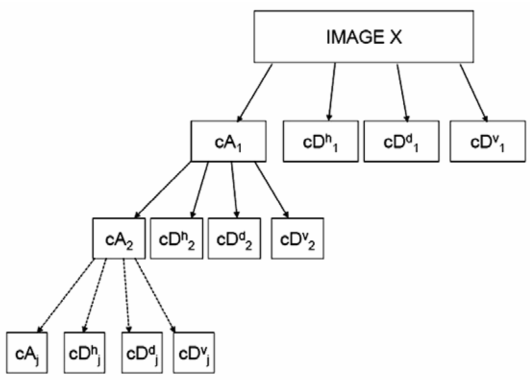


Fig. 2. Wavelet decomposition tree.

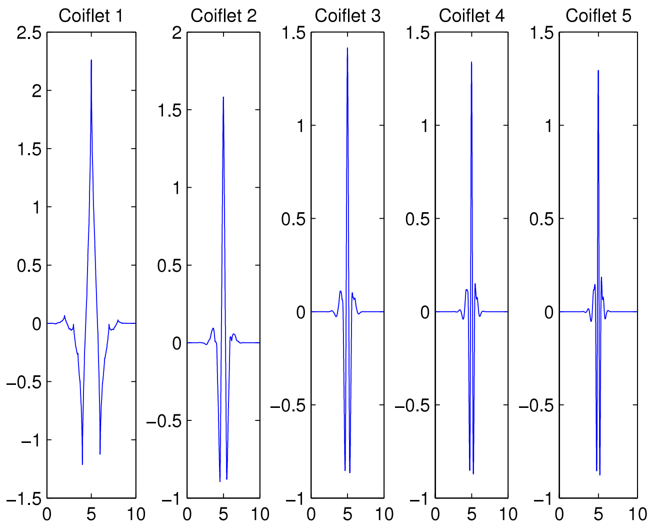


Fig. 3. Wavelet coiflets family.

$$\psi^1(\mathbf{x}) = \phi(x_1)\psi(x_2), \qquad \psi^2(\mathbf{x}) = \psi(x_1)\phi(x_2), \qquad \psi^3(\mathbf{x}) = \psi(x_1)\psi(x_2),$$

and denote for $1 \le i \le 3$

$$\psi^i_{j,\mathbf{k}}(\mathbf{x}) = \frac{1}{2^j}\psi^i\left(\frac{x_1 - 2^j i_1}{2^j}, \frac{x_2 - 2^j i_2}{2^j}\right).$$

The wavelet family $\{\psi^1_{j,i}, \ \psi^2_{j,i}, \ \psi^3_{j,i}\}_{i \in \mathbb{Z}^2}$ *is an orthonormal basis of* $\mathbf{W}^2_j$ *and* $\{\psi^1_{j,i}, \ \psi^2_{j,i}, \ \psi^3_{j,i}\}_{(j,i) \in \mathbb{Z}^3}$ *is an orthonormal basis of* $\mathbb{L}^2(\mathbb{R}^2)$.

The process of wavelet compression has the following stages to comment on:

1. Project the image in a subspace $\mathbf{V}_j$ with j sufficiently large.
2. Implement the algorithm of the wavelet decomposition. We performed a mathematical transformation on the object to achieve a sparse representation.
3. Establish a scheme of quantization of the processed data. If the coefficients in the decomposition are irrational numbers they must be rounded to the computer calculations.
4. Apply standard compression techniques to the object. These techniques allow us to compress the image in ratios of $20:1$ or higher ratios.
5. Perform the reverse process to decompress.

In this paper we want to show a comparison of different thresholds in order to select the best. For that, we do not use the additional methods of lossless compression to increase the compression ratio, such as Huffman coding or RLE (Run Length Encoding). We only treat the first and second stages. These stages contain three steps:

(a) Decomposition: a wavelet and a decomposition level j are chosen for computing the wavelet decomposition of the image X (Fig. 2).

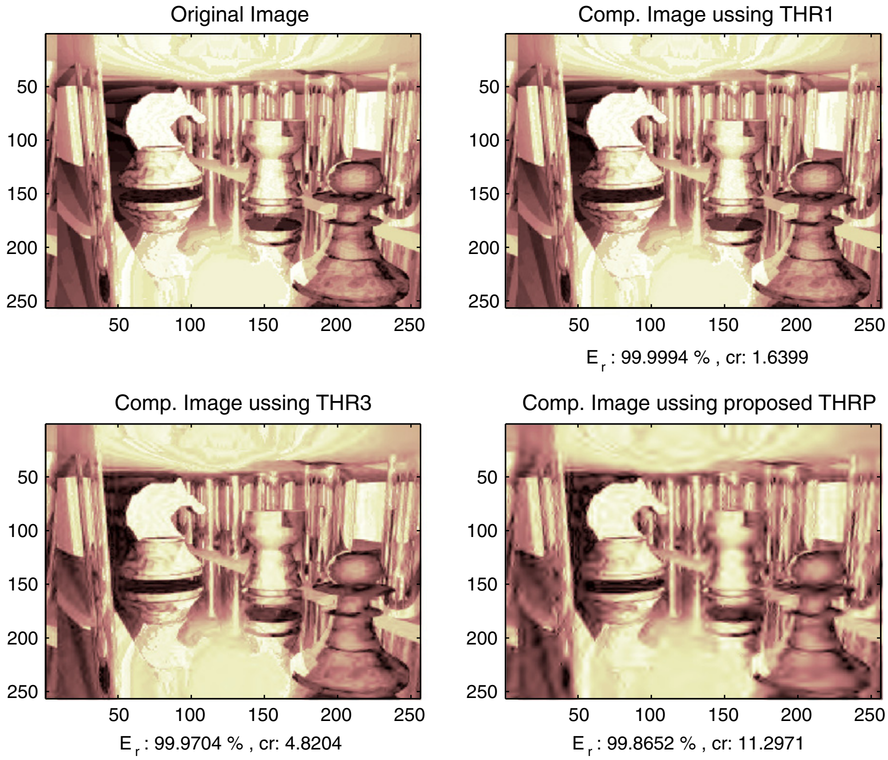


Fig. 4. Representation of the chess image and its compressed images using the thresholds *THR*1, *THR*3 and $\mathcal{THRP}$.

(b) Threshold detail coefficient: for each level from 1 to j, a threshold is selected and a hard thresholding function is applied to the detail coefficients, where the coefficients that do not exceed a certain threshold must be removed.
(c) Reconstruction: the wavelet reconstruction is computed using the original approximation coefficients of level j and the modified detail coefficients of levels from 1 to j.

Note that the reason for the loss of compression is the non-existence of the reverse in the thresholding step.

2. Methodology

We chose biorthogonal wavelets. These wavelets are generated by translations and dilations of two functions, one of which is used to analyze, and the other is used to reconstruct.

The mother wavelet family used is the 'coiflets' (coifN, $N = 1, \dots, 5$). It is an orthogonal wavelet family that is characterized by having a large number of zero moments ($2N$ for the wavelet function and $2N - 1$ for the scaling function); it is a biorthogonal and compactly supported (Fig. 3).

In order to obtain a good compression it is important to take into account the approximation coefficients and the decomposition level. The approximation coefficients save the most important features of the image and they can be used to reconstruct the image without some details. That is, they achieve a good compression ratio without a significant loss of quality in the signal. The decomposition level is important, because generally, if it is too low it does not get a good compression, while if the level is too high it can distort the image.

To analyze the efficiency of the compressor, we use as to parameters the energy retained E_r and the compression ratio cr given by

$$E_r = \frac{\|X_c\|^2}{\|X\|^2}100, \qquad pc0 = \frac{nz}{nc}100, \qquad cr = \frac{100}{100 - pc0},$$

where X_c represents the compressed image, X is the original image, nz is the number of zeros and nc is the number of coefficients.

In the compression we use a global threshold T. It is usual to employ these options:

(i) *THR*1, is a threshold close to 0. A typical THR value is the median of the absolute value of the detail coefficients of the first level.
(ii) *THR*2, is a threshold such that the percentages of the retained energy and the number of zeros are the same.
(iii) *THR*3, is a threshold equal to the square root of the *THR*2.

First, for a prefixed decomposition level we compare different wavelet coiflets for a same image (which size is 256×256 and 524 288 Bytes) and we obtain that the wavelet 'coif5' is the best for the different thresholds *THR*1, *THR*2 and *THR*3. Second, for this image and using this wavelet we compare different decomposition levels, and we conclude that $j = 5$ is the optimal.

Note that to assure the quality of the image, the loss of the energy after the compression technique must be less than 0.15%. Thus we eliminate *THR*2 since its retained energy is smaller than this quantity.

We propose a new global threshold $\mathcal{THRP}$ with a good *cr* and which $E_r \geq 99.85\%$. This threshold corresponds to a percentile P_{99} of the absolute value of the detail coefficients of the first level.

Once the decomposition level and the mother wavelet have been chosen, we compare the different global thresholds usually employed with the proposed threshold in several images.

3. Algorithm

We give the algorithm used in this paper applied in the Matlab program.

Step 1. Choose 'coif5' as mother wavelet.
Step 2. Choose $j = 5$ as level decomposition.
Step 3. Choose the thresholds *THR*1, *THR*3 and $\mathcal{THRP} = P_{99}$.
Step 4. Introduce $k = 99.85$.
Step 5. Calculations.
  Step 5.1. If $E_r(THR3) > k$ and $E_r(\mathcal{THRP}) > k$, then obtain E_r and *cr* of *THR*1, *THR*3 and $\mathcal{THRP}$.
  Step 5.2. If $E_r(THR3) < k$ or $E_r(\mathcal{THRP}) < k$ then $\mathcal{THRP} = P_t\ (/\ t \in [80, 99])$ is the value that makes $E_r \geq 99.85\%$ and obtain E_r and *cr* of *THR*1, *THR*3 and $\mathcal{THRP}$.
Step 6. Construct a table with the results obtained in *Step 5.1.* and *Step 5.2.*.
Step 7. Calculate the average of E_r and *cr*.

4. Results and conclusions

In this section we apply the proposed algorithm for 20 images; in 7 of them Step 5.2 has been used. In the following table we show the different parameters t used in these 7 images.

	t		t		t		t
facets	89	belmont2	86	geometry	84	tartan	89
wmandril	86	woman	97	detail1	85		

The results obtained using the algorithm are given in the following table.

	THR1 $E_r(\%)$	THR1 cr	THR3 $E_r(\%)$	THR3 cr	$\mathcal{THRP}$ $E_r(\%)$	$\mathcal{THRP}$ cr
belmont1	99.9983	1.7499	99.9484	6.4034	99.8816	13.2474
bust	99.9999	1.6189	99.9786	14.8610	99.9877	10.2655
chess	99.9994	1.6399	99.9704	4.8204	99.8652	11.2971
detfingr	99.9985	1.5253	99.9487	4.2415	99.9170	5.3566
mask	100.0000	1.5628	99.9918	20.4676	99.9976	12.6895
nbarb1	99.9973	1.6466	99.9529	5.0723	99.9051	8.2243
sinsin	100.0000	1.4190	99.9802	4.5238	99.9994	3.0358
tire	99.9996	1.5470	99.9525	6.3659	99.9211	8.0836
wbarb	99.9988	1.6747	99.9639	5.5598	99.8710	11.4317
cathe2	99.9999	1.6919	99.9621	12.3609	99.9596	12.8338
cathe1	99.9999	1.6895	99.9636	11.7699	99.9633	11.8282
wgatlin	99.9999	1.4537	99.9381	6.9002	99.9795	4.7828
woman2	99.9992	1.5706	99.9600	4.8691	99.9001	7.9607
facets	99.9998	1.9702	99.9201	4.7569	99.8538	5.8384
belmont2	99.9898	1.6790	99.9494	2.5162	99.8511	3.8861
geometry	99.9994	1.6226	99.8727	4.0968	99.8732	4.0925
tartan	99.9920	1.4498	99.9242	2.1596	99.8520	2.6201
wmandril	99.9861	1.6385	99.9143	2.7654	99.8600	3.6292
woman	99.9947	1.8790	99.9374	6.2364	99.8799	10.7730
detail1	99.9943	1.6338	99.9811	1.9443	99.8537	3.4998
	99.9973	1.6331	99.9505	6.6346	99.9086	7.7688

$THR1$ shows an average of the retained energy $E_r = 99.9973\%$ but its compression ratio (cr) is very small, of the order of $1.6 : 1$.

If we compare $THR3$ and $\mathcal{THRP}$ we note that the loss in E_r is 0.04%, while the improvement in cr is 1.13% and the image result is successful (Fig. 4).

We believe that the results are quite satisfactory. However, if we need more resolution in the images our proposed threshold cannot be adequate.

Acknowledgement

Supported by Spanish Grant MMT2007-64477.

References

[1] C.M. Brislawn, Fingerprints go digital, Notices of the AMS 42 (11) (1995) 1278–1283.
[2] I. Daubechies, Ten lectures on wavelets, in: CBMS-NSF Conference Series in Applied Mathematics, SIAM Ed., 1992.
[3] S. Mallat, W.L. Hwang, Singularity detection and processing with wavelets, IEEE Trans. Inform. Theory. 38 (1992) 617–643.
[4] S. Mallat, Wavelet Tour of Signal Processing, Academic Press, USA, 1999.