A Statistical Analysis of the 2D Discrete Wavelet Transform

Alexandru Isar[1], Sorin Moga[2], and Xavier Lurton[3]

[1] University POLITEHNICA
2 Bd. V. Parvan,
1900 Timisoara, Romania
(e-mail:isar@etc.utt.ro)
[2] GET - ENST Bretagne
Technopôle de Brest Iroise, CS 83818,
29238 BREST Cedex, France
(e-mail: sorin.moga@enst-bretagne.fr)
[3] IFREMER - Centre de Brest
Service Acoustique et Sismique (TMSI/AS)
BP 70 Plouzané 29820 France
(e-mail: lurton@ifremer.fr)

Abstract. The aim of this paper is a complete statistical analysis of the two dimensional discrete wavelet transform, 2D DWT. This analysis represents a generalization of the statistical analysis of the 1D DWT, already reported in literature. The probability density function, the correlation and the first two moments of the coefficients of the 2D-DWT are computed. The asymptotic behaviour of this transform is also studied. The results obtained were used to design a new denoising system dedicated to the processing of SONAR images.
Keywords: Discrete Wavelet Transform, Asymptotic analysis, convergence speed.

1 Introduction

The 2D DWT is a very modern mathematical tool. It is used in compression (JPEG 2000), denoising and watermarking applications. To exploit all its advantages, it must be carefully analyzed. The aim of this paper is the study of this transform from the statistical point of view. Such a complete study was not already reported.

2 The 2D DWT

In this paper the most commonly used 2D DWT is considered. It is built with separable orthogonal mother wavelets, having a given regularity. At every iteration of the DWT, the lines of the input image (obtained at the end of the previous iteration) are low-pass filtered with a filter having the impulse response m_0 and high-pass filtered with the filter m_1. Then the lines of the two images obtained at the output of the two filters are decimated with a factor of 2. Next, the columns of the two images obtained are low-pass filtered with m_0 and high-pass filtered with m_1. The columns of those four images are also decimated with a factor of 2. Four new sub-images (representing the result of the current iteration) are generated. The first one, obtained after two low-pass filterings, is named approximation sub-image (or LL image). The others three are named detail sub-images: LH, HL and HH. The LL image represents the input for the next iteration. In the following, the coefficients of the DWT will be noted with ${}_xD_m^k$, where x represents the image who's DWT is computed, m represents the iteration index (the resolution level) and $k = 1$, for the HH image, $k = 2$, for the HL image, $k = 3$, for the LH image and $k = 4$, for the LL image. These coefficients are computed using the following relation:

$$ {}_xD_m^k\left[n,p\right] = \left\langle x\left(\tau_1,\tau_2\right), \psi_{m,n,p}^k\left(\tau_1,\tau_2\right)\right\rangle \tag{1}$$

where the wavelets can be factorized:

$$\psi_{m,n,p}^k\left(\tau_1,\tau_2\right) = \alpha_{m,n,p}^k\left(\tau_1\right)\cdot\beta_{m,n,p}^k\left(\tau_2\right) \tag{2}$$

and the two factors can be computed using the scale function $\varphi\left(\tau\right)$ and the mother wavelets $\psi\left(\tau\right)$ with the aid of the following relations:

$$\alpha_{m,n,p}^k\left(\tau\right) = \begin{cases}\varphi_{m,n}\left(\tau\right), \; k=1,4\\ \psi_{m,n}\left(\tau\right), \; k=2,3\end{cases} \tag{3}$$

$$\beta_{m,n,p}^k\left(\tau\right) = \begin{cases}\varphi_{m,n}\left(\tau\right), \; k=2,4\\ \psi_{m,n}\left(\tau\right), \; k=1,3\end{cases} \tag{4}$$

where:

$$\varphi_{m,n}\left(\tau\right) = 2^{-\frac{m}{2}}\varphi(2^{-m}\tau - n) \tag{5}$$

$$\psi_{m,n}\left(\tau\right) = 2^{-\frac{m}{2}}\psi\left(2^{-m}\tau - n\right) \tag{6}$$

3 The pdfs of the wavelet coefficients

These pdfs can be computed following the description of the 2D DWT given in the previous paragraph. In fact each sub-image has its own pdf. The pdfs computation is based on the relation between the pdfs of the random variables from the input and the output of a digital filter. This is a sequence of convolutions which number is equal with the number of the filter coefficients. The pdfs of the wavelet coefficients, ${}_xD_m^k$, can be expressed with the aid of the pdf of the input image, x, using the relation, [1]:

$$f_{{}_xD_m^k}\left(a\right) = \star_{q_1=1}^{M(k)}...\star_{r_m=1}^{M_0} f_d(k,q_1,r_1,...,q_m,r_m,a) \tag{7}$$

where:

$$f_d\left(k, q_1, ..., r_m, a\right) = G\left(k, q_1, ..., r_m\right) f_x\left(G\left(k, q_1, ..., r_m\right) a\right) \tag{8}$$

and:

$$G\left(k, q_1, ..., r_m\right) = \frac{1}{F\left(k, q_1, r_1\right) \prod_{l=2}^{m} m_0\left[q_l\right] m_0\left[r_l\right]} \tag{9}$$

where:

$$F\left(k, q_1, r_1\right) = \begin{cases} m_0\left[q_1\right] m_0\left[r_1\right], & for \;\; k=4 \\ m_0\left[q_1\right] m_1\left[r_1\right], & for \;\; k=3 \\ m_1\left[q_1\right] m_0\left[r_1\right], & for \;\; k=2 \\ m_1\left[q_1\right] m_1\left[r_1\right], & for \;\; k=1 \end{cases} \tag{10}$$

M_0 represents the length of the impulse response $m_0, M_1$the length of m_1 and the numbers of the first two groups of convolutions in relation (7) are given by the relation:

$$M\left(k\right) = \begin{cases} M_0, & for \;\; k=4 \\ M_0, & for \;\; k=3 \\ M_1, & for \;\; k=2 \\ M_1, & for \;\; k=1 \end{cases} \; and \; N\left(k\right) = \begin{cases} M_0, & for \;\; k=4 \\ M_1, & for \;\; k=3 \\ M_0, & for \;\; k=2 \\ M_1, & for \;\; k=1 \end{cases} \tag{11}$$

In conformity with (7), each pdf of the wavelet coefficients is a sequence of convolutions. Hence, the random variable representing the wavelet coefficients can be written like a sum of independent random variables. So, the central limit theorem can be applied. This is the reason why the pdf of the wavelet coefficients tends asymptotically to a Gaussian, when the number of convolutions in (7) (the DWT iterations number) tends to infinity. This number depends on the mother wavelets used and on the number of iterations of the DWT. For mother wavelets with a long support, this number becomes large very fast (for a small number of iterations). The mother wavelet with the shortest support is the Haar mother wavelets. We have computed, using the relation (7), the pdfs of the coefficients of the 2D DWT of an image, containing a noise distributed following a $log - gamma$ distribution, using the Haar mother wavelets. The support of the mother wavelets used in practice is longer than the support of the Haar mother wavelets, considered in this theoretical case. The difference between the pdfs of the wavelet coefficients obtained after the second iteration and Gaussians is small in this case. So, after two iterations, the pdfs of the wavelet coefficients can be considered Gaussians. For the first two iterations, heavy-tailed models must be considered. Finer analysis, measuring the distance between the real pdfs and Gaussians, are performed in [Foucher *and al.*, 2001], [Achim *and al.*, 2003] and [Xie *and al.*, 2002].

4 The correlation of the wavelet coefficients

The input image, x, represents, in general, the sum of the useful image, s, and of the noise image, n. Because these two random signals are not correlated, the correlation of the wavelet coefficients of the image x, is the sum of the correlations of the wavelet coefficients of the useful image and of the noise image. The correlation function of the wavelet coefficients can be computed using the following relation:

$$\Gamma_{{}_xD_m^k}\left[n_1,n_2,p_1,p_2\right]=E\left\{{}_xD_m^k\left[n_1,p_1\right]\left({}_xD_m^k\left[n_2,p_2\right]\right)^*\right\}$$

$$=\int_{R^4}E\left\{x\left(\tau_1,\tau_2\right)\right\}\cdot\psi_{m,n_1,p_1}^{k*}\left(\tau_1,\tau_2\right)\cdot\psi_{m,n_2,p_2}^{k}\left(\tau_3,\tau_4\right)d\tau_1 d\tau_2 d\tau_3 d\tau_4 \qquad (12)$$

or:

$$\Gamma_{{}_xD_m^k}\left[n_1,n_2,p_1,p_2\right]=\frac{1}{4\pi^2}\int_{R^2}\gamma_x\left(2^{-m}\nu_1,2^{-m}\nu_2\right)\cdot$$

$$\cdot\left|\alpha_2\left\{\psi^k\left(\nu_1,\nu_2\right)\right\}\right|^2\cdot e^{-j\left[\nu_1\left(n_2-n_1\right)+\nu_2\left(p_2-p_1\right)\right]}d\nu_1 d\nu_2 \qquad (13)$$

where the first factor under the integral from the right hand side represents the power spectral density of the input image and the second factor represents the power spectral density of the one dimensional mother wavelets used. In the following, the influence of each of these two factors will be analyzed. For the beginning, the influence of the first factor is considered. If the input image is a white noise, with a known variance, z, it can be written:

$$\gamma_n\left(2^{-m}\nu_1,2^{-m}\nu_2\right)=z \qquad (14)$$

and the expression of the wavelet coefficients of the input noise image correlation function becomes:

$$\Gamma_{{}_nD_m^k}\left[n_1,p_1\right]=z\cdot\delta\left[n_1\right]\cdot\delta\left[p_1\right] \qquad (15)$$

This relation was obtained applying some very well known results from harmonic analysis: the Wiener-Hincin identity and the symmetry theorem. A magic property of the orthogonal wavelet bases (the samples of the correlation functions of the corresponding mother wavelets and scaling functions, taken at integer moments, are discrete-time unit impulses) was also used. Hence, the correlation of the wavelet coefficients of a white noise image do not depends on the regularity of the one dimensional mother wavelets used. The same result can be obtained taking in (13) the limit for m (number of iterations) tending to infinity. Indeed, under the integral from the right hand side of (13), only the power spectral density of the input image depends on m. After the limit computation, this function becomes a constant, like in the case when the input image is a white noise. Asymptotically, the 2D DWT transforms every colored noise into a white one. Hence this transform can be regarded as a whitening system, for any regularity of the one dimensional mother wavelets used. So, the wavelet coefficients sequences of the noise component of the input image are white noise sub-images, having the same variance. In the following, some considerations about the influence of the second factor of the product under the integral from the right hand side of the relation (13), will be made. This second factor takes into account the specific of the one dimensional mother wavelets used. It explains how the regularity of the wavelet decomposition affects the coefficients correlation. It can be proved that the convergence speed to a white noise (when m tends to infinity) increases when the regularity (the length of the filters m_0 and m_1) increases. So, the convergence speed to a Gaussian white noise can be increased using one dimensional mother wavelets with higher regularity. The first and second order moments of the wavelet coefficients can be computed using the following relations.

$$E\left\{{}_xD_m^k\left[n_1,p_1\right]\right\}=E\left\{\int_{R^2}x\left(\tau_1,\tau_2\right)\cdot\psi_{m,n_1,p_1}^{k*}\left(\tau_1,\tau_2\right)d\tau_1 d\tau_2\right\}= \tag{16}$$

$$=\begin{cases}0, k=1,2,3\\ 2^m\cdot\mu_x, k=4\end{cases}$$

Only the means of the images formed with the approximation wavelet coefficients are not nulls. The mean of the DWT of the noise component of the input image is given by the relation:

$$E\left\{{}_nD_m^k\left[n_1,p_1\right]\right\}=\begin{cases}0, k=1,2,3\\ -2^m\cdot\mu_n, k=4\end{cases} \tag{17}$$

In practice the number of iterations of the DWT is important. The dimensions of the image built with the approximation wavelet coefficients obtained after the last iteration are smalls. This is the reason why this image is not filtered in the denoising applications based on the use of the DWT. The variance of the wavelet coefficients of the noise component can be computed using the relation:

$$\sigma^2_{{}_xD_m^k}=E\left\{\left|{}_xD_m^k\left[n_1,p_1\right]\right|^2\right\}=\varGamma_{{}_xD_m^k}\left(0,0\right)=$$

$$=\frac{1}{4\pi^2}\int_{R^2}\gamma_x\left(2^m\nu_1,2^m\nu_2\right)\cdot\left|\alpha_2\left\{\psi^k\left(\nu_1,\nu_2\right)\right\}\right|^2 d\nu_1 d\nu_2$$

The DWT of the input noise component, n, has a variance given by:

$$\sigma^2_{{}_nD_m^k}=\begin{cases}z, & k=1,2,3\\ z-2^{2m}\mu_n^2, & k=4\end{cases} \tag{18}$$

This variance is constant for all the images formed using detail wavelet coefficients. Hence, it can be estimated using the first HH image. This estimation can be used for the filtering of any other detail image, formed with the detail wavelet coefficients obtained at any iteration. The correlation of the DWT of s is given by:

$$\Gamma_{{}_sD_m^k}[n_1,p_1] = 2^{2m}\cdot\Gamma_s[2^m n_1, 2^m p_1] \tag{19}$$

its mean by:

$$E\left\{{}_sD_m^k[n_1,p_1]\right\} = \begin{cases} 0, & k=1,2,3 \\ 2^m\cdot\mu_s, & k=4 \end{cases} \tag{20}$$

and its variance, by:

$$\sigma^2_{{}_sD_m^k} = 2^{2m}\cdot\sigma_s^2 \tag{21}$$

So, the variance of the detail wavelet coefficients sequences obtained starting from the useful component of the input image increases when the iteration index increases. All the relations established in this paragraph were used in [Isar and Moga, 2004], for the design of a denoising system for SONAR images.

5 Conclusion

A complete analysis of the 2D DWT was reported. It is proved that the 2D DWT asymptotically converges to the 2D Karhunen-Loève transform. So, the DWT of a colored noise image, with a given probability density function, converges asymptotically to a white Gaussian noise. This is a generalization of the results reported in [Isar *and al.*, 2002], where the case of the 1D DWT was considered. Another reference for the statistical analysis of the 1D DWT is [Pastor and Gay, 1995]. The asymptotic analyses of 1D DWT and 2D DWT have similar results. The pdfs of both wavelet transforms converge asymptotically to Gaussians. Both wavelet transforms converge asymptotically to the corresponding Karhunen-Loève transforms, for any regularity of the one dimensional mother wavelets used. The convergence speed to a Gaussian white noise can be improved increasing the regularity of the one dimensional mother wavelets used. Both wavelet transforms convert a white noise into a white noise with the same variance. All the other results of the statistical analyses of the 1D DWT and 2D DWT (pdfs, correlations, moments) are also similar. Based on the statistical analysis reported in this paper, a new denoising system was built in [Isar and Moga, 2004]. Its performances for the treatment of the SONAR images are also reported. This statistical analysis can be used for compression or watermarking purposes also. Statistical analyses of other wavelet transform will be reported soon.

References

[Foucher *and al.*, 2001]Samuel Foucher, Gozé Bertin Bénié, Jean-Marc Boucher, "Multiscale MAP Filtering of SAR images", *IEEE Transactions on Image Processing*, vol. 10, no.1, January 2001, 49-60.

[Achim *and al.*, 2003]Alin Achim, Panagiotis Tsakalides and Anastasios Bezerianos, "SAR Image Denoising via Bayesian Wavelet Shrinkage Based on Heavy-Tailed Modeling", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 8, August 2003, 1773-1784.

[Xie *and al.*, 2002]Hua Xie, Leland E. Pierce and Fawwaz T. Ulaby, "Statistical Properties of Logarithmically Transformed Speckle", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 3, March 2002, 721-727.

[Isar and Moga, 2004]A. Isar, S. Moga, Le débruitage des images SONAR en utilisant la transformée en ondelettes discrète à diversité enrichie, Rapport de recherche, LUSSI-TR-2004-4, Département Logiques des Usages, Sciences Sociales et Sciences de l'Information, Laboratoire Traitement Algorithmique et Matériel de la Communication, de l'Information et de la Connaissance, CNRS FRE 2658, ENST-Bretagne, 2004.

[Isar *and al.*, 2002]A. Isar, A. Cubitchi, M. Nafornita, Algorithmes et techniques de compression, Ed. Orizonturi Universitare Timisoara, 2002.

[Pastor and Gay, 1995]D. Pastor, R. Gay, "Décomposition d'un processus stationnaire du second ordre. Propriétés statistiques d'ordre 2 des coefficients d'ondelettes et localisation fréquentielle des paquets d'ondelettes", Traitement du signal, vol. 12, no. 5, pp. 393-420, 1995.