Digital Signal Processing 24 (2014) 95–105

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

On the statistical decorrelation of the 2D discrete wavelet transform coefficients of a wide sense stationary bivariate random process

A. Isar *, C. Nafornita

Communications Department, Politehnica University of Timisoara, Blvd. Vasile Parvan 2, 300223 Timisoara, Romania

ARTICLE INFO

Article history:
Available online 9 October 2013

Keywords:
Wavelet transforms
Correlation
Random process
White noise
Colored noise

ABSTRACT

We present a second order statistical analysis of the 2D Discrete Wavelet Transform (2D DWT) coefficients. The input images are considered as wide sense bivariate random processes. We derive closed form expressions for the wavelet coefficients' correlation functions in all possible scenarios: inter-scale and inter-band, inter-scale and intra-band, intra-scale and inter-band and intra-scale and intra-band. The particularization of the input process to the White Gaussian Noise (WGN) case is considered as well. A special attention is paid to the asymptotical analysis obtained by considering an infinite number of decomposition levels. Simulation results are also reported, confirming the theoretical results obtained. The equations derived, and especially the inter-scale and intra-band dependency of the 2D DWT coefficients, are useful for the design of different signal processing systems as for example image denoising algorithms. We show how to apply our theoretical results for designing state of the art denoising systems which exploit the 2D DWT.

1. Introduction

A great number of Wavelet Transforms (WT) as for example: 2D Discrete WT (2D DWT) [1], 2D Undecimated DWT (2D UDWT) [2], 2D Dual Tree Complex WT (2D DTCWT) [3], etc., can be used for image processing, because most of the image information is concentrated in few large wavelet coefficients, property known as sparsity of the wavelet representation.

This simplifies and accelerates the image processing algorithm considered and is a consequence of the 2D WT decorrelation properties.

The first results about the decorrelation effect of WT were obtained for 1D transforms. For example the covariance of coefficients obtained by wavelet decomposition of random processes can be computed recursively based on an algorithm described in [4]. This algorithm has an interesting link to the 2D DWT, which makes computations faster. A statistical analysis of 1D DWT was reported in [5] and it was generalized in [6] to the wavelet packets case. Some results of statistical analysis of 2D WT can also be found. In [7] is treated the case of 2D DWT, highlighting the inter-scale and inter-band dependencies of wavelet coefficients, with the aid of the mutual information concept, but closed form expressions for the correlation functions are missing. A statistical analysis of 2D UDWT is presented in [8] and a second order statistical analysis of 2D DTCWT is presented in [9].

* Corresponding author.
E-mail addresses: alexandru.isar@etc.upt.ro, isar1alexandru@yahoo.com (A. Isar), corina.nafornita@etc.upt.ro, corina.nafornita@gmail.com (C. Nafornita).

1051-2004/$ – see front matter
http://dx.doi.org/10.1016/j.dsp.2013.10.001

All the WT are characterized by two features: the mother wavelets (MW) and the primary resolution (PR), or the number of decomposition levels. The importance of their selection is highlighted in [10]. An appealing particularity of 2D DWT is the inter-scale dependency of the wavelet coefficients [7]. The goal of the present paper is a complete second order statistical analysis of the 2D DWT, establishing closed form expressions for the correlation functions in all four possible scenarios. We also highlight the influences of the 2D DWT features on that correlation functions.

Every image denoising method has three steps: acquired signal's WT computation; filtering in wavelets domain; computation of Inverse WT (IWT). A huge number of denoising methods were developed in the last years, by associating different WT with different filters (requested in the second step of the denoising method). In the following, we present a possible classification of denoising methods, highlighting each class with examples based on 2D DWT. A first category of denoising methods is composed by non-parametric techniques. These are denoising methods which not take into account any model of the components of the acquired signal [11]. A second category of denoising methods is composed by parametric techniques, [12–14], which consider statistical models for both components of the acquired image. Many of these methods are based on the utilization of Maximum A Posteriori (MAP) filters, in the second step. The construction of a MAP filter necessitates statistical models for the noiseless and noise components of the image to be filtered. Finally, there are some denoising methods, which lie at the border of parametric and non-parametric techniques, named semi-parametric techniques [15–18]. These consider models only for the noise component of the input image.

In conjunction with the expansion of new wavelet estimators, some researchers have worked on improving the wavelet transform itself. Since the early – non-redundant – 2D DWT, substantial improvements have been reached in denoising by using shift-invariant transformations, as the 2D UDWT [8,19], or quasi-shift invariant WT with better directional selectivity, as for example the 2D DTCWT [20–22], the steerable pyramid [23], the Dual-Tree *M*-band WT (DTMBWT) [24,25] or the Hyperanalytic WT (HWT) [12]. The new properties resulting from the use of often highly redundant transforms (as for example the 2D UDWT) have been obtained at the expense of the loss of orthogonality, a substantially more intensive memory usage and a higher computational cost than that of the 2D DWT. The latter point becomes a major concern in image volume denoising and more generally in multichannel image denoising, in particular when the number of channels is large. For instance, even though the usual color image representations require no more than 3–4 channels (RGB, HSV, YUV, or CMYK descriptions), the computational cost is already quite large when shift-invariant (i.e., undecimated) transforms are involved.

The structure of this paper is the following. In the second section we study the statistical decorrelation of the 2D DWT coefficients when the image is a wide sense stationary bivariate random process, developing the results presented in [26]. Starting from the implementation of this transform, we highlight the four possible scenarios: inter-scale and inter-band, inter-scale and intra-band, intra-scale and inter-band and intra-scale and intra-band dependencies. We treat the case of the 2D DWT coefficients of a bivariate white Gaussian noise (WGN) as well. The most important theoretical results obtained in the second section are verified by simulation in the third section, where some experimental results are presented. The object of the fourth section is a discussion of the results presented in previous sections, oriented toward image denoising. Finally, the conclusions of the paper are presented in the fifth section.

2. A second order statistical analysis of 2D DWT

The main advantage of 2D DWT versus other 2D WT, as for example the 2D DTCWT, is its computational flexibility, as it inherits some of the classes of MW developed in the framework of the 1D DWT, like the Daubechies, Symmlet or Coiflet families [27]. This non-redundant transform can be implemented using the very fast Mallat's algorithm [1]. The drawbacks of the 2D DWT are lack of translation invariance and poor directional selectivity. The perfect translation invariance can be reached using the 2D UDWT. Quasi-translation invariance can be obtained using Complex WT (CWT) as for example the 2D DTCWT or the HWT [12]. It represents a natural generalization of the 2D DWT, which is conceived for real images, for hyperanalytic images. The lack of translation invariance of 2D DWT can be corrected in denoising application [15] (see Section 4). Both CWT already mentioned have also better directional selectivity than 2D DWT.

2.1. 2D DWT implementation

Each of the iterations of the Mallat's algorithm implies several operations [1]. The rows of the input image, obtained at the end of the previous iteration, are passed through two different filters: a low-pass filter – L with the impulse response m_0 and a high-pass filter – H with the impulse response m_1, resulting in two different sub-images. The rows of the two sub-images obtained at the output of the two filters are decimated with a factor of two. Next, the columns of the two images obtained are filtered with m_0 and m_1. The columns of those four sub-images are also decimated with a factor of two. Four new sub-images, representing the result of the current iteration (which corresponds to the current decomposition level – or scale), are obtained. These sub-images are called subbands. The first sub-image, obtained after two low-pass filtering (LL), is named approximation sub-image (or LL subband). The other three are named detail sub-images: LH, HL and HH. The LL subband represents the input for the next iteration. In the following, the coefficients of 2D DWT will be denoted by d_m^k, where m represents the current scale and k is the subband and it is $k=1$ – for LH, $k=2$ – for HL, $k=3$ – for HH and $k=4$ – for LL. These coefficients are computed using the following scalar products:

$$d_m^k[n,p]=\langle f(\tau_1,\tau_2),\psi_{m,n,p}^k(\tau_1,\tau_2)\rangle,\quad \tau_1\in R,\ \tau_2\in R, \tag{1}$$

where f represents the image whose 2D DWT is computed (considered as a bivariate random process) and the wavelets are real functions and can be factorized as:

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

and the two factors can be computed using the scaling function $\varphi(\tau)$ and the MW $\psi(\tau)$ with:

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

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

where:

$$\varphi_{m,n}(\tau)=2^{-\frac{m}{2}}\varphi\big(2^{-m}\tau-n\big)\quad\text{and}$$
$$\psi_{m,n}(\tau)=2^{-\frac{m}{2}}\psi\big(2^{-m}\tau-n\big). \tag{5}$$

Taking into account Eqs. (3)–(5), it can be written:

$$\psi_{m,n,p}^k(\tau_1,\tau_2)=2^{-m}\psi^k\big(2^{-m}\tau_1-n,2^{-m}\tau_2-p\big),$$
$$\text{where } \psi^k(\tau_1,\tau_2)=\psi_{0,0,0}^k(\tau_1,\tau_2). \tag{6}$$

This is the case of the so-called separable MW. Bivariate MW which do not satisfy Eq. (2), called non-separable, exist as well, but are less popular, because the algorithms of the corresponding 2D DWT are more complicated and slower. They are not considered in the present paper.

2.2. The expectation of the wavelet coefficients

We begin the second order statistical analysis by computing the statistical mean of the wavelet coefficients:

$$\mu_{d_m^k}=E\{d_m^k\}=E\{\langle f(\tau_1,\tau_2),\psi_{m,n,p}^k(\tau_1,\tau_2)\rangle\}$$
$$=E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\tau_1,\tau_2)\psi_{m,n,p}^{k*}(\tau_1,\tau_2)\,d\tau_1\,d\tau_2\right\}. \tag{7}$$

Applying Fubini's theorem and taking into account the fact that the random process f is wide sense stationary, we obtain:

$$\mu_{d_m^k}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}E\{f(\tau_1,\tau_2)\}\psi_{m,n,p}^{k*}(\tau_1,\tau_2)\,d\tau_1\,d\tau_2$$
$$=\mu_f\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_{m,n,p}^{k*}(\tau_1,\tau_2)\,d\tau_1\,d\tau_2$$
$$=\mu_f\cdot\mathrm{F}^*\{\psi_{m,n,p}^k(\tau_1,\tau_2)\}(0,0). \tag{8}$$