# Cyclostationary error analysis and filter properties in a 3D wavelet coding framework

Riccardo Leonardi, Alberto Signoroni *

*Department of Electronics for Automation - University of Brescia, Italy*
*via Branze, 38 I-25123 Brescia, Italy*

**Abstract**

The reconstruction error due to quantization of wavelet subbands can be modelled as a cyclostationary process because of the linear periodically shift variant property of the inverse wavelet transform. For audio and image coding applications this fact is of little practical interest, both from quantitative and perceptual points of view. Differently, when coding is applied to volumetric data (such as MR scans) or sequences of images (such as video) the non stationary nature of the reconstruction error becomes an objectionable concern. Indeed as 3D data are usually displayed as series of images, the reconstruction error is typically measured by the $PSNR[n]$ sequence, where $n$ represents either a spatial slicing plane (3D data) or the temporal reference plane (video). What can be observed is an oscillation of $PSNR[n]$ typically of the order of 1dB. This work aims at founding an explanation to this phenomenon. The analysis is entirely carried out in the original signal domain and can easily be extended to more than 3 dimensions. In the case of biorthogonal wavelet basis it is shown that a substantial reduction of the oscillation amplitude can be achieved under a proper selection of the basis functions. Different coding rates have been considered from high resolution conditions to low bit-rate coding. An experimental validation of the proposed model has been performed in the case of a bit-plane coding approach.

*Key words:* Wavelet coding, 3D data, video coding, linear phase biorthogonal filters, cyclostationary error model, bit-plane quantization

* Corresponding author: phone +39 030 3715 432, fax +39 030 380014
*Email addresses:* riccardo.leonardi@ing.unibs.it (Riccardo Leonardi), alberto.signoroni@ing.unibs.it (Alberto Signoroni).

# 1 Introduction

## *1.1 Overview and objectives*

The quantization error introduced in the wavelet domain is often simply considered to be equivalent to a quantization performed in the signal domain due to the (near-) orthogonal behavior of the discrete wavelet transform (DWT). While this is reasonably true with respect to the error variance, high order dependencies of the quantization error in the signal domain may lead to adverse effects for certain types of signals. The linear and periodically time-variant (LPTV) nature of the inverse DWT modifies the statistical properties of the quantization error process introduced in the transformed domain. In terms of stationarity the error becomes cyclostationary (CS) [1]. In this work, the effects of quantization error non linear dependencies are considered for the case of multidimensional data coding. The CS statistical behavior will be associated to the data sampling grid and characterized by a hypercubic periodic pattern characterized by the time-varying periodicity along each data dimension. Statistical behavior will be measured on hyperplanes intersecting the data, e.g. the $PSNR$ measured on the rows of an image or on the slices of a volume. In these cases, our interest is focusing on how and to what extent the CS property generates oscillatory phenomena of the PSNR by moving the hyperplane along the remaining orthogonal directions. This point is not commonly considered in image coding: no particular attention is dedicated to fluctuations of the PSNR between consecutive rows, as the perception of the coded information remains 2-dimensional. In the case of the slices of a volume instead, measuring the $PSNR[n]$ on each slice $n$ orthogonally to the slicing plane is a common procedure. Indeed in this case only minimal fluctuations can be tolerated in particular in the case of medical data-sets. As the slicing can take place in any particular direction, the oscillatory behavior should be monitored along all such directions. Objective quality fluctuations are usually tolerated in the case of (2D+t) video compression (where they are typically generated by hybrid schemes) because of the band-pass temporal response of the human visual system. On the contrary, when considering static 3D data (e.g. computed tomography CT or magnetic resonance MR data-sets) adjacent slices are commonly seen side by side or within a movie loop, with low or hand-driven frame-rate. In this case limiting $PSNR$ oscillation is important to ensure good reliability with respect to the original material. This is especially true in the field of diagnostic imaging, where 3D images are in widespread use and compression is a relevant concern. This paper proposes a quantitative analysis and a multidimensional model in order to describe the reconstruction error oscillation due to wavelet based compression and how to ease its reduction. In fact a particular property of some linear phase biorthogonal filters will emerge which causes a great reduction of the oscillation. A time domain

formulation of the problem is proposed here so as to describe the pattern of oscillation. Section 2 addresses the problem from one to $L$ wavelet decomposition levels in 1D. Section 3 extends the 1D case to the multidimensional one (especially 3D). The presentation is then adapted to handle the characteristics of real codecs working at various bit-rates (section 4). Experimental simulation are described finally in section 5.

### *1.2 Relations to other works*

The periodically time variant property of the wavelet transform is well known [1]. The analysis of the effects of this property in terms of propagation of the subband quantization error on the reconstructed data has been made by Uzun and Haddad in [2] as an extension of the work of Westerink *et al.* [3]. The main objective of this work was to model and find wavelet filter banks which minimize the reconstruction error in an $MSE$ sense. Those approaches, and in general other works on subband coding systems which incorporates quantization models (e.g. [4–7]), highlight expected values of the spectral density functions and perform their analysis mainly in the z-transformed domain. In a similar framework, Reichel *et al.* [8] compare the subband quantization effects for real and integer wavelet transforms. In this work, a different perspective has been adopted: we are interested in modelling not the expected value but the non stationary properties of the reconstruction error in the data-domain. As a consequence a data-domain analysis – which allows to assess the relationship between local statistical behavior and filter-bank properties – is suggested.

As far as the coding schemes are concerned, a certain number of papers describe effective wavelet compression scheme for 3D data [9–14], while a standardization activity is currently being finalized by the ISO/JPEG community [15]. In the validation phase of this work we do not consider a specific coding scheme. This is due to the fact that for most existing coders similar bit-plane quantization strategies are adopted.

## 2 Reconstruction error properties

In this section we shortly review some well known fundamentals of multirate filter banks and we start to use a one-dimensional framework to formulate a data domain statistical analysis of the reconstruction error. We consider first a single stage 2-channel decomposition, then extend to the case of a multilevel DWT. Finally, the focus is shifted to the case of linear phase biorthogonal wavelet basis. In order to easily characterize the error terms, the polyphase

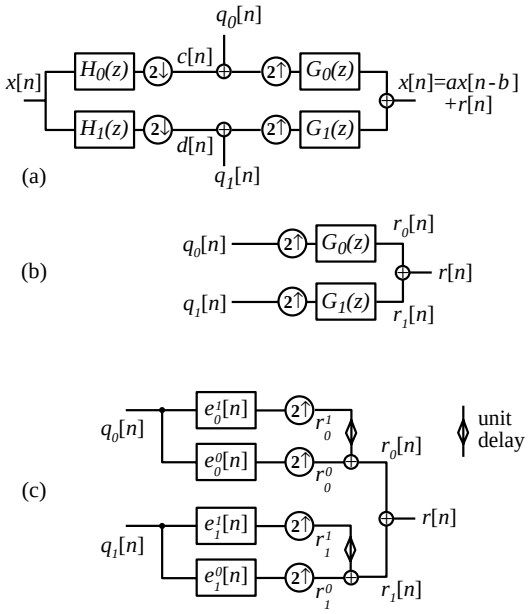


Fig. 1. *(a) 2-channel subband decomposition and quantization error addition; (b) Synthesis of the reconstruction error and its equivalent polyphase representation (c).*

multirate filter bank representation (see e.g. [1]) is a more effective representation, with respect to the octave-band tree structure.

### *2.1 Reconstruction error in a single synthesis stage*

The additive quantization error, introduced with respect to the $i$-th subband as denoted by the signals $q_i[k]$. At the output of the inverse discrete wavelet transform (IDWT), the *reconstruction error* $r[n]$ adds up to the perfect reconstructed signal. Each subband error $q_i[k]$ contribution to $r[n]$ is denoted by $r_i[n]$. The statistical properties of $q_i[k]$ and the linear periodically time variant (LPTV) nature of each channel of the synthesis filter bank determine a CS behavior of each $r_i[n]$, thus of the whole $r[n]$. Fig.1(a) shows the well known 2-channel decomposition with the associated two sources of quantization noise in the subband domain. Without such noise components $q_i[n]$, the overall system acts as a pure delay $b$ with gain $a$. The quantization error contribution shown in Fig.1(b) creates a LPTV [1] behavior, with period 2, due to the presence of the dyadic up-samplers. The polyphase representation of the synthesis bank is shown in Fig.1(c), where

[1] From now on we prefer to use the term "shift-variance" instead of "time-variance" because of the nature of the data that will be considered. For the sake of clarity the acronym LPTV will be preserved.

$$
\begin{aligned}
g_i[n] &= (2\uparrow)\left(e_i^0\right)[n] + (2\uparrow)\left(e_i^1\right)[n-1]\,, \\
G_i(z) &= E_i^0(z^2) + z^{-1}E_i^1(z^2)\,.
\end{aligned}
\tag{1}
$$

Here various error components [2] can be considered for their contribution to the reconstruction error $r[n]$, namely $r = r_0 + r_1 = r_0^0 + r_0^1 + r_1^0 + r_1^1 = (r_0^0 + r_1^0) + (r_0^1 + r_1^1) = r^0 + r^1$, where $r^0 \triangleq r_0^0 + r_1^0$ and $r^1 \triangleq r_0^1 + r_1^1$ represent the overall contribution of the polyphase components at even and odd samples respectively, that is $r^0[n] = 0$ for $(n \,\mathsf{m}od 2) = 1$ and $r^1[n] = 0$ for $(n \,\mathsf{m}od 2) = 0$.

To better understand the synthesis filter bank effect on the quantization error we start from a high resolution analysis: for each channel $i$, $q_i[n]$ can be considered the realization of a uniformly distributed ergodic white noise $\mathbf{q}_i[n]$. In addition, $\forall i, \bar{n}$ let us assume that the random variables $\mathbf{q}_i[\bar{n}]$ are i.i.d., with $f_{\mathbf{q}_i}(\alpha) = \frac{1}{\Delta} rect(\frac{\alpha}{\Delta})$. Hence they have the same expected value and variance $\eta_q = E\{q_i[n]\} = 0$ and $\sigma_q^2 = E\{q_i^2[n]\} = \frac{\Delta^2}{12}$.

In the simple case of Fig.1(c) it's easy to verify that the 2-dimensional ergodic input processes lead, thanks to the LPTV property, to a one-dimensional cyclostationary output process, with period 2, having the following characteristics: by linearity, $\eta_r = 0$ ; by the LPTV properties,

$$
\sigma_r^2[n] = \begin{cases} \sigma_{r^0}^2 & \text{for} \quad (n \,\mathsf{m}od\, 2) = 0\,, \\ \sigma_{r^1}^2 & \text{for} \quad (n \,\mathsf{m}od\, 2) = 1\,. \end{cases}
$$

Now, always by linearity, the reconstruction error can be decomposed as follows:

$$
\sigma_{r^0}^2 = \sigma_{r_0^0}^2 + \sigma_{r_1^0}^2 = \sigma_{q_0}^2 \mathcal{G}^2\left(e_0^0[n]\right) + \sigma_{q_1}^2 \mathcal{G}^2\left(e_1^0[n]\right) \tag{2}
$$

$$
\sigma_{r^1}^2 = \sigma_{r_0^1}^2 + \sigma_{r_1^1}^2 = \sigma_{q_0}^2 \mathcal{G}^2\left(e_0^1[n]\right) + \sigma_{q_1}^2 \mathcal{G}^2\left(e_1^1[n]\right)\,, \tag{3}
$$

where $\mathcal{G}^2(\cdot)$ is the energy operator and $e_j^i[n]$ represents each filter polyphase component.

[2] We adopt a notation in which the superscripts correspond to the reference instants in the period of time variance (or the reference number of the polyphase components), while subscripts are used to identify each subband channel. Moreover, we indicate as $(2\uparrow)(\cdot)$ the upsampling operator on its argument.

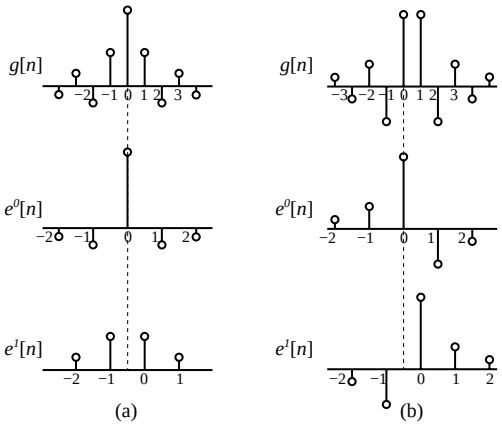


Fig. 2. *Polyphase components of a linear phase FIR filter in the case of (a) even length, (b) odd length.*

### *2.2 Reduction of the variance oscillation*

The absolute difference $|\sigma^2_{r^0} - \sigma^2_{r^1}|$ is a first indicator of the entity of the cyclostationary implications on the reconstruction error. As it can be intuitive it is desirable to make this difference small. Considering (2) and (3), and supposing that $\sigma^2_{q0} = \sigma^2_{q1}$, for the above difference to be equal to zero it is sufficient that the sum of the energies of the low and high pass corresponding (same phase) polyphase components give the same value. Minimization of the $|\sigma^2_{r^0} - \sigma^2_{r^1}|$ could be conversely achieved by proper design of the filter bank. As it will be shown, a very simple solution can be found thanks to the properties of linear phase biorthogonal filters. These filters are commonly used for visual data compression. They have either odd (e.g. the popular 9/7 filters [16]) and even lengths (e.g. the 10/18 filers introduced in [17] and the more recent 22/14 ones [18]). It is well known that linear phase filters exhibit symmetric or antisymmetric impulse response. Let us now understand how this property is conveyed to the polyphase representation of a linear phase filter bank. Odd length symmetric filters (Type I and III linear-phase FIR) have sample centered symmetry axes, while even length ones (Type II and IV) have bin centered symmetry axes. Thus it is easy to notice (see Fig.2) that in the odd length case the polyphase splitting generates two symmetric filters (an odd length and an even length one), while for even length filters two non-symmetric filters are generated, one being the mirror image of the other (with respect to the original center of symmetry). At this point, we observe that for even length filters the cyclostationary behavior of the output noise $r[n]$ for a 2-channel filter bank, is structurally eliminated because $\sigma^2_{r^0} = \sigma^2_{r^1}$ and $f_{\mathbf{r}^0}(\alpha) = f_{\mathbf{r}^1}(\alpha)$, whereas this does not hold for the odd length filter case. Even though this clear-cut distinction is only valid on a 1-D wavelet decomposition case, it appears instructive for other filterbank configurations.

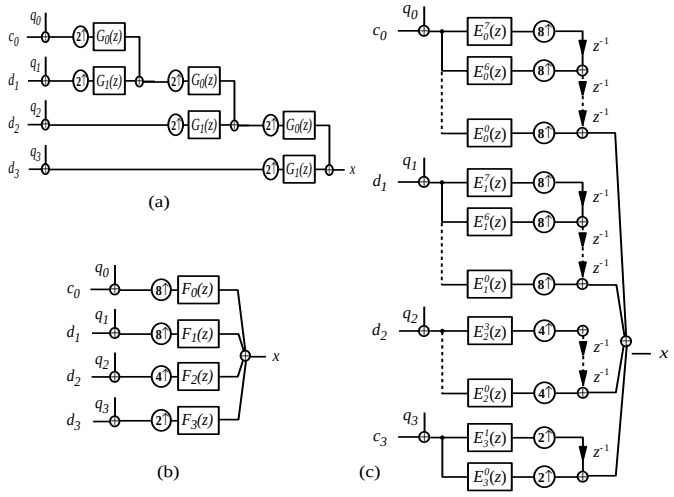


Fig. 3. *A 3 level IDWT: (a) octave band (dyadic) subdivision tree, (b) single-adder multirate filter bank, (c) polyphase representation of (b).*

### *2.3 Reconstruction error in a L-levels IDWT*

The analysis of the output noise characteristics is now considered for a complete octave band wavelet decomposition. In Fig.3a a dyadic $L$-level IDWT is shown (with $L = 3$). For the purpose of establishing the statistical behavior of the output noise component, it is convenient to use the equivalent single-adder $(L+1)$-channel multirate scheme (Fig.3b), where $F_0(z) = G_0(z^4)G_0(z^2)G_0(z)$, $F_1(z) = G_1(z^4)G_0(z^2)G_0(z)$, $F_2(z) = G_1(z^2)G_0(z)$ and $F_3(z) = G_1(z)$. The polyphase representation (Fig.3c) allows for an immediate interpretation of the reconstruction error components.

The value of $\sigma_r^2$ at the position $n$ can be obtained by a generalization of (2) and (3) which considers a higher number of subband channels with the appropriate number of polyphase components:

$$\sigma_r^2[n] = \sum_{l=0}^{L} \sigma_{r_l}^2[n] = \sum_{l=0}^{L} \sigma_{r_l^{n \bmod \lambda(l)}}^2 \quad , \quad \lambda(l) = 2^{L+1-(l+\delta[l])} \tag{4}$$

being $\sigma_{r_l^j}^2 = \sigma_{q_l}^2 \cdot \mathcal{G}^2\left(e_l^j[m]\right)$, with $l = 0, \ldots, L$ and $j = 0, \ldots, \lambda(l) - 1$, the $j$-th polyphase component of $\sigma_{r_l}^2[n]$. In (4) we see that at each position $n$ the right polyphase component $j = n \bmod \lambda(l)$ is selected for each subband channel $l$. The usage of $\delta(l)$ allows to handle any low pass channel 0 with the same resolution level of the high pass channel 1.

The random process r$[n]$ which represents the reconstruction error is cyclostationary with period $2^L$, in fact r$[n]$ represents a sum of CS random processes r$_l[n]$ with periodicity $2^l$. We show in Fig.4 the sequence of $\sigma_{r_l}^2[n]$ values for a 16 sample signal, with $L$ set to 3 and with a common value for each subband

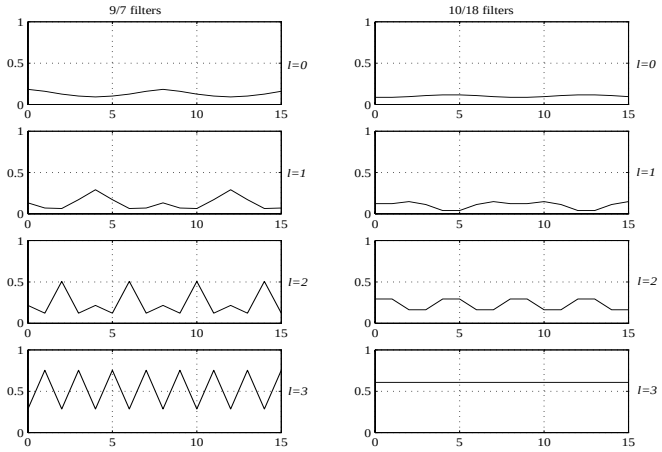


Fig. 4. *CS $\sigma^2_{r_l}[n]$ oscillation magnitude (l: resolution level) r 9/7 filters (a), and 10/18 ones (b) for a 16-sample signal and a 3 level IDWT.*

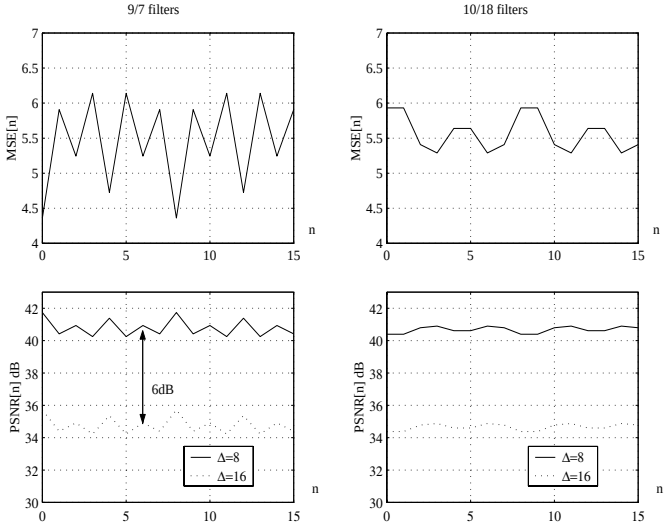


Fig. 5. *Overall $MSE$ and $P_{SNR}$ fluctuations for a 3 level 1D-IDWT.*

error variance $\sigma_q^2 = \Delta^2/12$ with $\Delta = 8$. During the analysis stage the lower resolution subbands exhibit a larger gain with respect to the high resolution ones. This gain is compensated during the synthesis stage so that the wavelet domain quantization error is subject to a de-emphasis which becomes more substantial as the resolution level decreases. This can be seen from Fig.4 where the average error contribution decreases towards low resolution subbands and so does its oscillation amplitude. Therefore the first decomposition level error contribution will be dominant in terms of defining the oscillation trend of the CS behavior. For identical biorthogonal wavelet filters, we plot in Fig.5 $MSE[n] = \sigma_r^2[n]$ and the corresponding $P_{SNR}[n]$ for each position $n$ in the original signal domain. The CS oscillation pattern exhibits as expected a $2^L$ $(L = 3)$ periodicity. The $P_{SNR}$ difference between the $j$-th and the $(j + 1)$-th

position for a $L$-level IDWT can be written as

$$\Delta P_{SNR}[j] = 10\log\frac{255^2}{\sigma^2_{r^j}} - 10\log\frac{255^2}{\sigma^2_{r^{j+1}}} = 10\log\frac{\sigma^2_{r^{j+1}}}{\sigma^2_{r^j}} \tag{5}$$

where the $j$-th $MSE$ value can be expressed as $\sigma^2_{r^j} = \sum_{l=0}^{L}\sigma^2_{r_l^j}$ , with $l = 0,\ldots,L$, $\sigma^2_{r_l^j}$ being the contribution of the $l$-th subband.

We can also note that for a reference quantization noise $\sigma_q^2$ which remains the same in each subband, there is no dependency of the $P_{SNR}$ fluctuation with respect to it. $\Delta P_{SNR}[j]$ is only a function of the polyphase filter coefficients; e.g. for a 1-level IDWT case (see (2),(3)):

$$\begin{aligned}|\Delta P_{SNR}[j]| = &\,10\log\left[\mathcal{G}^2\left(e_0^1[n]\right) + \mathcal{G}^2\left(e_1^1[n]\right)\right] \\ &- 10\log\left[\mathcal{G}^2\left(e_0^0[n]\right) + \mathcal{G}^2\left(e_1^0[n]\right)\right] \;\forall j.\end{aligned}$$

As expected the oscillation amplitude is smaller in the case of even length filters.

## 3 Three dimensional extension of the model

The above formulation provides the mean to model the reconstruction error in the context of $N$-dimensional data coding. In this section we concentrate on a set of structural properties of r[$n$] in a 2D or 3D framework. We make the assumption of dealing with separable wavelet transforms. This is common to almost all coding applications. Moreover, in order to avoid burdening the notation, we will consider only the case of same basis filters used for all spatial directions.

### *3.1 2D and 3D cyclostationarity patterns*

For N-dimensional data, periodic behavior of the reconstruction error statistics can be observed on the signal cartesian reference system. In particular the CS periodicities produce some elementary patterns (e.g. tiles or brick) which in turns partition the whole data space. Our first objective is to model what happens when increasing the problem dimension. Let us deal first with the 2D case. To calculate the error reconstruction variance at each pixel position the separable 2D IDWT dyadic tree is implemented using a separable single-adder representation. Dyadic and single-adder schemes are depicted in Fig.6 considering a 3 level decomposition. For each subband channel in Fig.6(b)

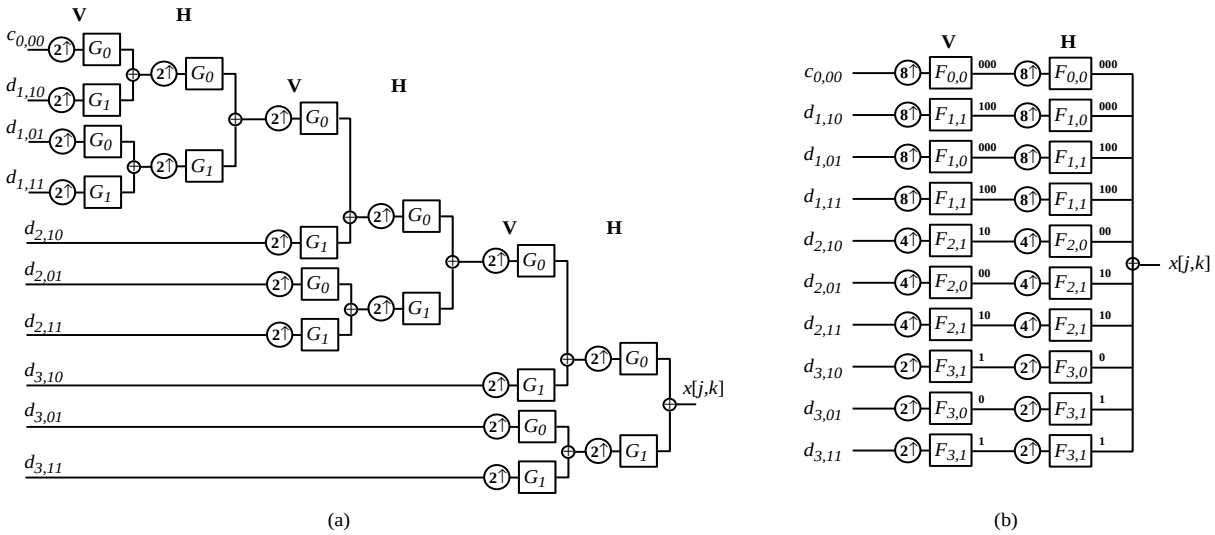


Fig. 6. *Dyadic (a) and single-adder (b) representation of a 3 level 2D-IDWT.*

two one-dimensional filters are operating in cascade along the vertical and horizontal dimensions respectively. A general expression of such filters, for $l = 1 \dots L$, is given by

$$F_{l,a}(z) = \prod_{i=l}^{L} G_{a \cdot \delta[i-l]}(z^{\gamma(i)}) \tag{6}$$

with $a \in \mathbb{B} = \{0, 1\}$, $\gamma(i) = 2^{L-i}$ . For $l = 0$, $F_{0,0}(z) = F_{1,0}(z)$ (because the $c_0$ and $d_1$ coefficients belong to the same scale). The filters involved in the horizontal and vertical directions are assumed to be the same (then, in (6), we don't use direction related subscript for the $z$ variable). In the first synthesis stage of Fig.6(b), the vertical filtering introduces on each subband column the CS behavior described by the (2), (3). It can be easily observed that, at the input of the horizontal stage (before the upsamplers), i.i.d. hypotheses on the error still hold along the row direction, but with alternating non-uniform pdf's on each row due to the previous column filtering. The horizontal stage introduces a CS behavior on each row, which combined to that introduced on the column, results in a square tiled bidimensional CS pattern. This simple mechanism is shown in Fig.7 in the case of 1-level IDWT (which determines a $2 \times 2$ cyclostationary pattern). By analogy to the 1D case, increasing the IDWT number of levels to $L$, the CS pattern becomes a $2^L \times 2^L$ matrix which constitutes the basic tile of the 2D cyclostationarity. In other terms, considering the reconstruction error variance, we have $\sigma_r^2[n_0, n_1] = \sigma_r^2[\bar{n}_0, \bar{n}_1]$ where $\bar{n}_0 = n_0 \bmod 2^L$, $\bar{n}_1 = n_1 \bmod 2^L$.

The reconstruction error variance values for each spatial (tile) position can be

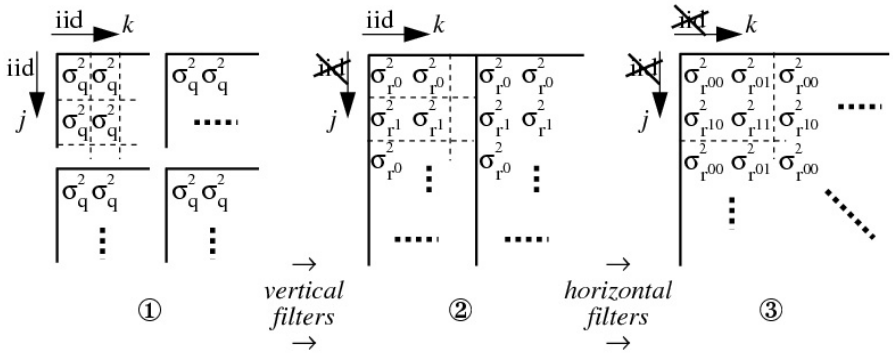


Fig. 7. 2D ciclostationary pattern generation for a 1-level IDWT.

calculated by the following extension of (4):

$$\sigma_r^2[n_0, n_1] = \sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{0,(0,0)}} + \sum_{l=1}^{L} \sum_{s\in\mathbb{B}^{2+}} \sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l,s}} \tag{7}$$

where $\bar{n}_0 = n_0 \bmod \lambda(l)$, $\bar{n}_1 = n_1 \bmod \lambda(l)$, $\lambda(l) = 2^{L+1-(l+\delta[l])}$ and, given $a, b \in \mathbb{B} = \{0, 1\}$,

$$\mathbb{B}^{2+} = \{(a, b)\} \setminus \{(0, 0)\}\,.$$

In (7) the first term refers to the lowest resolution ($l = 0$) low-pass subband coefficients ($s = (0, 0)$), while the second term represents the contribution of the 3 detail subbands $s$ at each resolution level $l \in (1, L)$. For each CS tile position, every subband contribution $\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l,s}}$ to the total reconstruction error variance is given by

$$\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l,s}} = \sigma^2_{q_{l,s}} P^{\bar{n}_0}_{l,a} P^{\bar{n}_1}_{l,b} \tag{8}$$

where $P^{\bar{n}}_{l,a} = \mathcal{G}^2\left(e^{\bar{n}}_{l,a}[n]\right)$ and $E^{\bar{n}}_{l,a}(z)$ the $\bar{n}-$th polyphase component of $F_{l,a}(z)$ defined in (6). In (8) we considered the general case of a different quantization error variance $\sigma^2_{q_{l,s}}$ associated to each subband.

A symmetry property can be demonstrated for the $\sigma_r^2[\bar{n}_0, \bar{n}_1]$ CS tile pattern which is relevant for our analysis:

**Proposition 1** *Given $\bar{\bar{\sigma}}_r^2$ a $2^L \times 2^L$ matrix (CS tile) with elements $\sigma_r^2[\bar{n}_0, \bar{n}_1]$ derived from (7), then $\bar{\bar{\sigma}}_r^2$ is symmetric, i.e. $\bar{\bar{\sigma}}_r^2 = \left(\bar{\bar{\sigma}}_r^2\right)^T$.*

*Proof*: We easily observe that, thanks to the separability in (8), and to the complementarity of the elements of $\mathbb{B}^{2+}$, we have a symmetric condition already at each IDWT decomposition level, namely $\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{0,(0,0)}} = \sigma^2_{r^{\bar{n}_1,\bar{n}_0}_{0,(0,0)}}$ and $\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l}} = \sigma^2_{r^{\bar{n}_1,\bar{n}_0}_{l}}$ , since

$$\begin{aligned}\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l}} &= \sum_{s\in\mathbb{B}^{2+}} \sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l,s}} \\ &= \sigma^2_{q_{l,s}} \left(P^{\bar{n}_0}_{l,0} P^{\bar{n}_1}_{l,1} + P^{\bar{n}_0}_{l,1} P^{\bar{n}_1}_{l,0} + P^{\bar{n}_0}_{l,1} P^{\bar{n}_1}_{l,1}\right)\end{aligned}$$

even if, in general, the single subband component of the reconstruction error variance is not symmetric: $\sigma^2_{r^{\bar{n}_0,\bar{n}_1}_{l,s}} \neq \sigma^2_{r^{\bar{n}_1,\bar{n}_0}_{l,s}}$. Then, by denoting a fundamental CS component with the use of $2^{l+\delta[l]} \times 2^{l+\delta[l]}$ dimensional matrices $\bar{\bar{\sigma}}^2_{r_l}$ with elements $\sigma^2_{r^{n_0,n_1}_l}$, such matrices are symmetric by construction. Now (7) can be written in a matrix form: $\bar{\bar{\sigma}}^2_r = \bar{\bar{\mathrm{O}}}^L \otimes \bar{\bar{\sigma}}^2_{r_{0,(0,0)}} + \sum_{l=1}^{L} \bar{\bar{\mathrm{O}}}^l \otimes \bar{\bar{\sigma}}^2_{r_l}$, where $\otimes$ is the Kronecker product, while $\bar{\bar{\mathrm{O}}}^l$ is a $\left(2^{(l+\delta[l])-1} \times 2^{(l+\delta[l])-1}\right)$ matrix containing all ones.
As the operator $\otimes$ preserves the symmetry and both $\bar{\bar{\sigma}}^2_{r_{0,(0,0)}}$ and $\bar{\bar{\sigma}}^2_{r_l}$ are symmetric, $\bar{\bar{\sigma}}^2_r$ is symmetric being a sum of symmetric matrices. □

The 3D extension follows easily. Using the same notation we have:

$$\sigma^2_r[n_0,n_1,n_2] = \sigma^2_r[\bar{n}_0,\bar{n}_1,\bar{n}_2] = \sigma^2_{r^{\bar{n}_0,\bar{n}_1,\bar{n}_2}_{0,(0,0,0)}} + \sum_{l=1}^{L} \sum_{s\in\mathbb{B}^{3+}} \sigma^2_{r^{\bar{n}_0,\bar{n}_1,\bar{n}_2}_{l,s}} \ , \tag{9}$$

with

$$\sigma^2_{r^{\bar{n}_0,\bar{n}_1,\bar{n}_2}_{l,s}} = \sigma^2_{q_{l,s}} P^{\bar{n}_0}_{l,a} P^{\bar{n}_1}_{l,b} P^{\bar{n}_2}_{l,c} \quad , \quad s=(a,b,c) \in \mathbb{B}^{3+} \ ; \tag{10}$$

and the following proposition holds:

**Proposition 2** *Given $\bar{\bar{\bar{\sigma}}}^2_r$ , a $2^L \times 2^L \times 2^L$ three-dimensional array (CS brick), with elements $\sigma^2_r[\bar{n}_0,\bar{n}_1,\bar{n}_2]$ derived from (9), then $\bar{\bar{\bar{\sigma}}}^2_r$ is symmetric, i.e. $\bar{\bar{\bar{\sigma}}}^2_r = \left(\bar{\bar{\bar{\sigma}}}^2_r\right)^{T_1} = \left(\bar{\bar{\bar{\sigma}}}^2_r\right)^{T_2}$, where $(\cdot)^{T_1}$ and $(\cdot)^{T_2}$ are the two possible transposition operators with respect to the main diagonals of the tensor; i.e. $\bar{\bar{\bar{\mathrm{A}}}} = (((\bar{\bar{\bar{\mathrm{A}}}})^{T_1})^{T_1})^{T_1} = (((\bar{\bar{\bar{\mathrm{A}}}})^{T_2})^{T_2})^{T_2}$ and $\bar{\bar{\bar{\mathrm{A}}}} = ((\bar{\bar{\bar{\mathrm{A}}}})^{T_1})^{T_2} = ((\bar{\bar{\bar{\mathrm{A}}}})^{T_2})^{T_1}$.* □

The proof of the this proposition can be derived in analog fashion to the proof used for Proposition 1. The analysis can easily be extended in a similar way to higher dimension problems.

Propositions 1 and 2 virtually allow to establish the spatial orientation invariance of the CS patterns which in turns determines, in the 3D case, the independence of the $P_{SNR}$ oscillation from the slicing direction, as we will see in the next section.

### *3.2 $P_{SNR}$ oscillation along various volume slicing directions*

The description of the 3D cyclostationarity on a symmetric cubic pattern allows us to quantify the fluctuations of the $P_{SNR}$ evaluated on the images obtained by slicing a volume (or 3D data-set) along one of its perpendicular axes. In fact, we can estimate the expected $MSE$ values along a certain slicing direction by averaging the punctual (voxel related) $MSE$ estimates of the

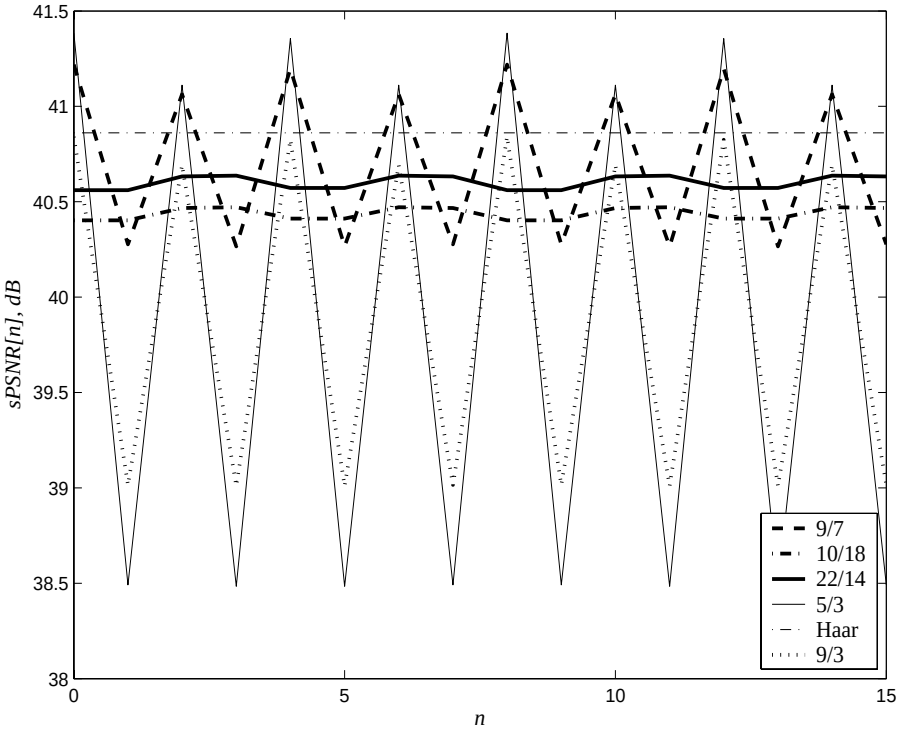


Fig. 8. *Quantification of the $sP_{SNR}[n]$ on adjacent cutting planes using expression (13). The direction n represents any of the 3 orthogonal directions of the volumetric data set. Six biorthogonal filters are compared, for a 3 level wavelet reconstruction.*

various CS brick slices. We call this measure "slice-based" $MSE$ ($sMSE$). For example, by slicing along the $n_0$ direction, we obtain the $2^L$ periodical sequence:

$$sMSE_r^0[n] = \frac{1}{2^{2L}} \sum_{n_1=0}^{2^L-1} \sum_{n_2=0}^{2^L-1} \sigma_r^2[n, n_1, n_2] \tag{11}$$

Thanks to the Proposition 2, we can easily observe that

$$sMSE_r^0[n] = sMSE_r^1[n] = sMSE_r^2[n] \quad \forall\, n\ , \tag{12}$$

where $sMSE_r^1[n] = \frac{1}{2^{2L}} \sum_{n_0=0}^{2^L-1} \sum_{n_2=0}^{2^L-1} \sigma_r^2[n_0, n, n_2]$,
and $sMSE_r^2[n] = \frac{1}{2^{2L}} \sum_{n_0=0}^{2^L-1} \sum_{n_1=0}^{2^L-1} \sigma_r^2[n_0, n_1, n]$.
Similarly we define a "slice-based" Peak Signal to Noise Ratio ($sP_{SNR}$) as

$$sP_{SNR}[n] = 10 \log \frac{255^2}{sMSE_r[n]}\ . \tag{13}$$

In Fig.8 we show the estimated $sP_{SNR}[n]$ values using (13) for 6 different biorthogonal filters, where we assumed to have the same quantization error $\sigma_{q_{l,s}}^2 = \sigma^2$ on each subband. The set of biorthogonal filters is representative of the most relevant wavelet kernels used in coding applications. It contains three odd length and three even length filter pairs. We included

- the best known B-spline 9/7 filter bank [16,19] and, from the same family, the 9/3 filters [19, pag. 277]
- two longer even length kernels, the 10/18 of [17] and the 22/14 [18]; both have demonstrated, also in our experience [12,20], competitive performance, despite their length, when used for coding 2D images and 3D medical data
- two well known short filters: the Haar and the 5/3 B-spline kernels [19, pag. 277], because of their role in the current research on wavelet video coding. In fact short filters are useful for an effective Motion Compensated Temporal Filtering (MCTF), which must be used instead of the rectilinear third dimension filtering, for high performance scalable video coding applications [21, Ch. 13].

Fig.8 suggests us to make the following considerations:

- As expected, odd length filters produce a substantially higher $sP_{SNR}[n]$ fluctuation with respect to the even length ones.
- The odd length filters 5/3 and 9/3, with their short (3 tap) low-pass synthesis stage, are particularly critical; they present a $sP_{SNR}[n]$ oscillation of about 2.8dB and 1.8dB respectively.
- Comparing the most popular filters for still image coding, the maximum $sP_{SNR}[n]$ oscillation for the 9/7 filters is about 0.95dB versus a 0.07dB of the 10/18 and 22/14 ones. In terms of reconstruction error power this means that we have a relative peak to peak $sMSE[n]$ oscillation (with respect to the mean value) of about 25% for 9/7 filters while it is only of the order of 1,5% for the 10/18 and 22/14 configurations.
- Filters do not only differs in terms of $sP_{SNR}[n]$ oscillation amplitude but also in terms of average $sP_{SNR}[n]$ value, and this fact can find a broader confirmation for N-dimensional signals. In other words we observe a filter ranking in terms of averaged quantization-reconstruction error power transfer. This could be seen as a contributing factor to the multivariate problem of determining the goodness of a wavelet basis for coding purposes. Usually the coding performance of wavelet filters are complex to quantify on real data and justifications are usually made in terms of approximation power or energy compaction attributes. Data centered criterions could be used in combination to quantization error centered ones (as the above ranking) in order to better understand the filter peculiarities for visual data coding. For example the Haar basis shows the lowest error power transfer, but it has poor approximation capabilities. On the other hand it is interesting to see that even if 22/14 and 10/18 visual performance should be hard to compare there clearly is an error power transfer disparity in favor of the 22/14 filter pair. Other considerations about the apparent better average behavior of the 9/7 filters will be made in the experimental section 5.
- Compared to the 1D case, there is an increased predominance of the finest detail subband on the overall oscillation trend (see Fig.4). This is due to the normalization (de-emphasis) factors applied to the lower detail subbands.

This originates from an augmented error de-emphasis (due to subband energy normalization) effect, which grows in power with the dimensionality of the problem.

- The plane based measure of $sP_{SNR}[n]$ actually averages the local expected $P_{SNR}[n_0, n_1, n_2]$ value on each voxel (or frame pixel). Thus the $sP_{SNR}$ oscillation between adjacent planes is less than the local $P_{SNR}$ pattern variations in the vicinity of a voxel. This observation further suggests to use even length linear phase filters also to keep a greater $P_{SNR}$ or $MSE$ homogeneity in a spatial neighborhood.

The above considerations will be further argued in Sec.5 by comparing the estimates deriving from (11) and shown in Fig.8 with respect to simulations on real data for a suitable rate-distortion range. Before to do this we want to describe what happens when, due to higher compression ratio, the oscillation model based on high resolution hypothesis, degrades.

## 4 3D embedded wavelet coding and bit-plane quantization

Until now we have made some simplifying hypothesis on $\sigma^2_{q_{l,s}}$, where $s \in \mathbb{B}^t$, $t \in \{1, 2, 3\}$. Our prediction of a CS fluctuation of the reconstruction error has to be reconsidered in a real context of interest. As it will be clear, the i.i.d. hypothesis has to be verified and discussed for various bit-rates from an intra- and inter-subband point of view. The properties of the quantization error are determined by the coding technique at various bit-rates and by the data to be coded which consequently affect the statistics of the wavelet subbands. Quantization optimality is well known for gaussian sources and high-resolution conditions [22–24]. At the same time, real data statistics and coding rate of interest are often far from such ideal conditions, and the associated behavior is no longer valid. In this section we present a degradation model which is able to describe what happens to the CS reconstruction error when diminishing the coding rate from high resolution to low bit-rate conditions. Due to the variability of the image statistics this will be a qualitative model that can be applied to a wide class of bit-plane based wavelet coders. Our specific interest consists in modelling the quantization error pdf when the bit-rate diminishes in order to infer on the value of $\sigma^2_{q_{l,s}}$. This in turns influences the $\sigma^2_r[\mathbf{n}]$ and $sP_{SNR}[n]$ fluctuations. To do this we briefly recall some essential prerequisites and assumptions before discussing the degradation model which will be verified in the experimental section.

### *4.1 Preliminary considerations on* $\sigma^2_{q_{l,s}}$

State of the art wavelet coders (based on zerotree [25,26], zeroblocks [27], significant-clusters [28,12] or independent blocks [29]), all make use of a progressive bit-plane quantization (BPQ) strategy. The wavelet structure and the high-order statistical dependency among the coefficients intra- or inter-subband are exploited by the significance map description and/or by the arithmetic coding to obtain a progressive or scalable coded bit-stream. The bit-plane quantizer is an almost uniform quantizer with the only exception of a double sized zero centered "dead zone". The corresponding quantization law can be expressed as:

$$Q(x,\Delta) = \begin{cases} 0 & \text{if } |x| < \Delta \ , \\ sgn(x)\left(\lfloor x/\Delta \rfloor + 1/2\right)\Delta & \text{if } |x| \geqslant \Delta \ ; \end{cases} \tag{14}$$

where, for the $b$-th bit-plane, $\Delta = 2^b$. When applied to the whole subband structure the progressive BPQ starts from the level $\Delta = 2^B$, with

$$B = \left\lfloor \max_{\mathbf{n},l,s} \left\{ \log_2 c_0[\mathbf{n}],\ \ \log_2 d_{l,s}[\mathbf{n}] \right\} \right\rfloor \tag{15}$$

and refines the wavelet coefficients representation by halving the value of $\Delta$ for each added bit-plane. Mallat [30] verified that at low bit-rates and for natural images the BPQ improves the R-D performance compared to the uniform one. Because of the zero concentrated wavelet coefficients pdf shape, a $2\Delta$ step size around the zero value considerably reduces the number of significant coefficients and consequently the associated bit-rate. The expected distortion is smaller than the one obtained by uniform $\Delta$ quantization with the same bit-rate reduction.

When departing from high resolution conditions the quantization error will become dependent of the data pdf. Pdf's of subband coefficients obtained from natural images have been observed to follow, with a good approximation, a generalized Gaussian distribution [16]:

$$\mathfrak{G}(x) = a \exp\left\{ - |bx|^{\gamma} \right\} \tag{16}$$

with

$$a = \frac{b\gamma}{2\Gamma\left(\frac{1}{\gamma}\right)} \quad \text{and} \quad b = \frac{1}{\sigma_x}\sqrt{\frac{\Gamma\left(\frac{3}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)}} \tag{17}$$

where $\Gamma(\cdot)$ is the "Gamma function"[31]. Given the source pdf (16) and the quantization law, it is well known how to precisely determine the quantization error pdf. Let $f_X(x)$ be the pdf of a subband source, and $\{I_k,\ k \in 1 \ldots K\}$ the set of events of the type $I_k : X \in (x_{k-1}, x_k]$ with probability $P_k = Prob\{I_k\}$,

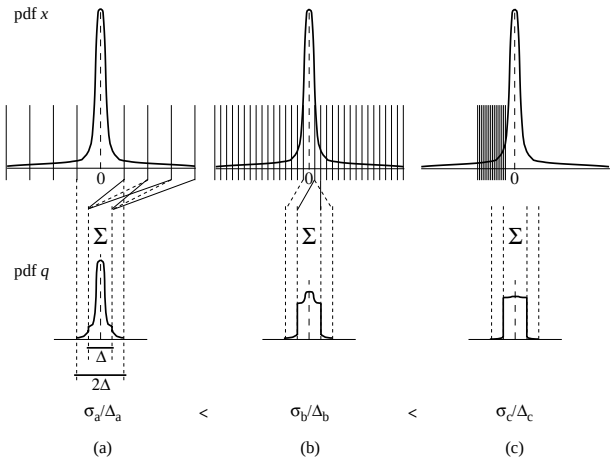


Fig. 9. *Assessment of the bit-plane quantization error pdf, starting from the original wavelet coefficient pdf. (a), (b) and (c) represents three different situation with increasing quantization resolution vs pdf coefficient variance ratio.*

then the quantization error pdf expression is given by (total probability theorem):

$$f_q(q) = \sum_{k=1}^{K} P_k f_q(q|I_k) \tag{18}$$

where the $f_q(q|I_k)$ is the zero-centered version of $f_X(x|I_k)$ with respect to the $I_k$ center value $x_k$,

$$f_q(q|I_k) = f_X(q + x_k|I_k) \tag{19}$$

with

$$f_X(x|I_k) = \begin{cases} f_X(x)/P_k & I_k \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{20}$$

For high-resolution conditions we can assume $f_q(q) = 1/\Delta \cdot rect(\Delta)$ with a good confidence. When high resolution cannot be assumed the original data pdf must be taken into account. The distribution described by (16), for $\gamma$ less than unity, exhibit a sharp peak around the origin. Then, if the quantizer belongs to the *midtread* or *dead-zone* class, this distribution peak falls, averaged as in (18) with the other levels, into the $f_q(q)$; on the contrary, with *midrise* quantizers the $f_q(q)$ remains uniform by construction, in fact, each $f_q(q|I_k)$ has its symmetrical counterpart. In addition, when using dead-zone quantizers, the non uniformity of the zero bin leads the $f_q(q)$ support to be higher than the $\Delta$ quantization step. In the BPQ case the maximum error is twice the uniform quantizer error with same $\Delta$: $|q|_{max} = \Delta$. However, $|q| > \Delta/2$ only for near zero coefficients and this is reassuring from a coding and artifact perception point of view. In Fig.9 we illustrate the above discussion, the $f_X(x)$ is processed using (18),(19) in three different situations. The $f_q(q)$ shape depends on the ratio of the $\Delta$ resolution with respect to the $\gamma$-modulated $f_X(x)$ shape. In fact, we consider three different conditions (a), (b) and (c). These may represent

alternatively a random variable with pdf $f_X(x)$ and quantized with $\Delta_a$, $\Delta_b$ and $\Delta_c$, or three different RV's pdf (16) with $(\sigma_{x_a}, \gamma_a)$, $(\sigma_{x_b}, \gamma_b)$, $(\sigma_{x_c}, \gamma_c)$ quantized with the same $\Delta$. Moreover, these three RV's could also be associated to 3 subbands of the same wavelet transform, or from another point of view, to the same subband for three different input data. We are interested to observe that, defining a *shape factor* $\varphi = \sigma_q/\Delta$ ($\varphi = 1/\sqrt{12}$ for uniform distributions), referring to Fig.9, $\varphi_a < \varphi_b < \varphi_c$ , and that in (8) as well as in (10) we can write:

$$\sigma^2_{q_{l,s}} = \varphi^2_{\Delta,l,s} \cdot \Delta^2 \ , \tag{21}$$

where

$$\varphi_{\Delta,l,s} = \varphi(\Delta, (\sigma_{x_{l,s}}, \gamma_{l,s})). \tag{22}$$

For a certain data to be coded, eq.(21-22) show the double dependence of the subband $\sigma^2_{q_{l,s}}$ from $l, s$ and the quantization interval $\Delta$ (which could be different e.g. for various subband and related to the coding rate or distortion).

### *4.2 Reconstruction error modelling for a bit-plane based wavelet coding*

We now reconsider the limitations introduced by the i.i.d. hypothesis (see Sec.2.1) on the subband based quantization error variances $\sigma^2_{q_{l,s}}$ which act in our CS model (see eqs (8), (10)). In particular we weaken the independence assumption and observe what happens to the subband based identical distribution (i.d.) from high to low bit-rates.

Statistical independence is only an ideal assumption. In order to model the output reconstruction error as a weighted sum of the subband quantization error (e.g. see (2) and (3)), it is sufficient to have uncorrelated subband quantization error samples. Thus the original hypothesis can be reduced to a wide sense CS (WSCS). Under the well known conditions of sufficient signal dynamics, low oversampling rate and lack of periodicity, a uniform quantizer produces a highly uncorrelated error even when using few quantization levels [22]. Due to the properties of the wavelet transform, the subbands contain critically sampled coefficients with a low degree of correlation, $\rho_x(n) \to 0$ for $n \neq 0$. However, at low bit rates a great amount of near zero coefficients falls into the dead zone of the BPQ. This leads to the evidence that most quantization error samples coincide with the near zero wavelet coefficients. This fact does not pose great problems in terms of dependencies as the wavelet coefficients are highly uncorrelated, but even if a "zero valued coefficient" correlation were present, its contribution to the overall correlation $\rho_q(1)$ remains moderate. Thus it is reasonable to consider uncorrelated quantization error between the different subbands.

Let us now track how the overall reconstruction error model changes as a function of the bit rate. For the sake of clarity we track the model degradations

on a discrete set of points each one reflecting a certain quantization bit-plane $b \in [1, B]$. For a given bit plane, the associated quantization step will be indicated as $\Delta_b$ which is applied to all the coefficients of all the subbands. As it can be expected for progressive or fine grain scalable coders, the transition between two adjacent bit-planes descriptions can be considered gradual (thus not explicitly described here). $B$ corresponds to the least significant bit-plane and thus $\Delta_B$ is the finest quantization step used by the coder. This represents the initial condition from which one can study how the model degrades. If $B$ is sufficiently large, one can assume that the high resolution i.i.d. conditions are valid for the whole subband set [3].

In such high resolution case $\varphi_{l,s}$ is worth about $1/\sqrt{12}$ and it becomes easy to calculate the reconstruction error properties, e.g. by using (9)-(13). Specifically, by considering two adjacent bit-plane, i.e. $\Delta_{b+1} = \Delta_b/2$, one can easily verify the classical 6dB difference e.g. between the $sP_{SNR}[n]$ levels along any slicing direction:

$$sP_{SNR}[n, b+1] = sP_{SNR}[n, b] + 6.02dB \ . \tag{23}$$

The 6dB value relies on the fact that, with a fixed shape-factor (21), all the $\sigma^2_{q_{l,s}}$ values are scaled by the same value as $\Delta_b$ is halved from bit-plane $b$ to bit-plane $b+1$.

This situation changes by decreasing the $b$, when a certain set of subbands modify their quantization error shape factor $\varphi$ value (see eq.(22)). In particular it is possible to observe that:

$$\varphi^2(\Delta_b, l, s) \leq \varphi^2(\Delta_{b+1}, l, s) \tag{24}$$

$$\varphi^2(\Delta_b, l, s) \geq \varphi^2(\Delta_b, l+1, s) \tag{25}$$

Equality holds when both terms of (24) or (25) are in high resolution conditions, while when these condition degrades, a twofold trend is described. Our qualitative observation relates principally to natural data (images or volumes [10]) and to the considerations made in Sec.4.1 about the subband and quantization error pdf models and the observations made on Fig.9. The relation (24) actually determines a shrinking of the average $sP_{SNR}$ distance among adjacent bit-planes, with respect to the 6dB rule of (23). The trend persists until the very last bit-planes (with $b$ near to 1) when another degeneration occurs for extremely low bit-rates, i.e. when the reconstructed data is useless. From inequality (25) the main effect is to introduce a certain decaying gradient of the shape-factor linked to an increase of the decomposition level $l$. Looking at (10) and (21) it is possible to understand how (25) determines a

[3] In the next section we will show that this is not always guaranteed and the quantization model could yet be degraded at the finest quantization level

smaller contribution from the high-level detail subbands to the CS oscillation pattern of the reconstruction error variance matrix $\bar{\bar{\sigma}}_r^2$, with respect to the case of high resolution conditions (in this latter case, see Sec.2.3 and Fig.4, the contribution of the high level subbands is dominant on the CS pattern). In short, looking to the $sP_{SNR}$ fluctuation of (13), the presence of (25) influences its amplitude and causes a reduction whose entity can be experimentally evaluated as shown in the next section.

## 5 Experimental observations and results

In the following some experiments are proposed in order to evaluate on real data the predictions of the reconstruction error model presented in the previous sections. A first part is deserved to the $sP_{SNR}$ fluctuation analysis with respect to the wavelet filter properties. Then, we discuss and show the visual impact of the $sP_{SNR}$ oscillations. Finally we verify the oscillation model degradation discussed in Sec.4. The 3D data we used are mainly medical volumetric data sets, coming from MR and CT 3D scanners. Visual results on a test video sequence will be presented too. The quality measures are extracted from single observations (intra data set) by exploiting the large number of voxels. In fact, thanks to a statistically reliable number of samples contained in a volume slice and to the ergodicity of the single cyclostationary components (cycloergodicity) of the cube, we can compare our statistical analysis with a single realization analysis. In particular, estimated values of $sMSE_r^i[n]$ can be compared with the measured $smse_r^i[n, x, \hat{x}]$, where for example:

$$smse_r^0[n, x, \hat{x}] = \frac{1}{D_1 D_2} \sum_{n_1=0}^{D_1-1} \sum_{n_2=0}^{D_2-1} (x[n, n_1, n_2] - \hat{x}[n, n_1, n_2])^2 \qquad (26)$$

with $D_{i=0,1,2}$ the size of the volume in voxels.

If not differently indicated, the progressive bit-plane based coding algorithm used for the following experiments is the the 3D version of our EMDC (Embedded Morphological Dilation Coding) algorithm [12,**?**].

### *5.1 Error fluctuations and filter properties*

The main objective of the following experiment is to verify the fluctuation model of Sec.3 and highlight the differences between even and odd length biorthogonal wavelet filters. For the test we selected the 9/7 and the 10/18 filter banks due to their widely recognized good coding performances. In Fig.10(a)-(d) we show the $sP_{SNR}$ measures associated to the considered filters

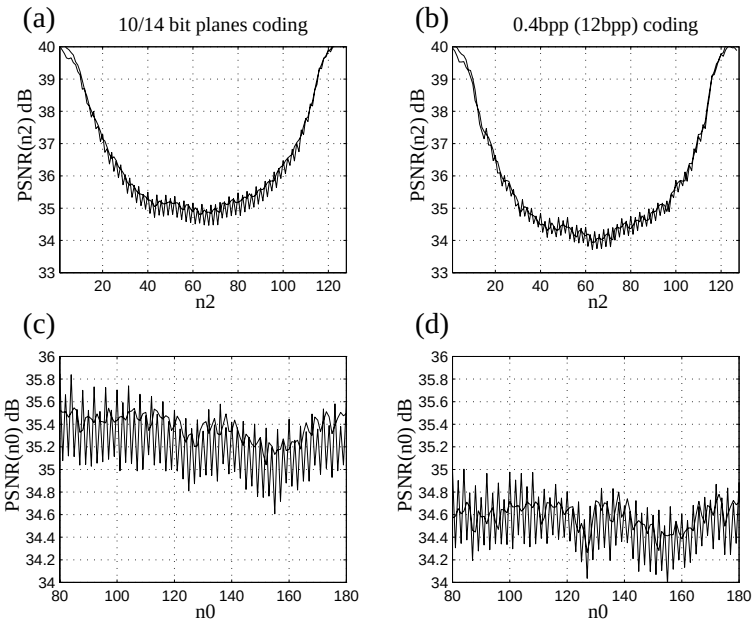


Fig. 10. *Comparison between coding with 9/7 filters (widest extent oscillation) and 10/18 ones, for the (256x256)x128 slice* MR-BRAINSAG *volume. In (a) and (b) the $sP_{SNR}$ curves are measured along the scanning axis; in (c) and (d) a windowed portion of the same curves is measured along the $n_o$ axis. (a) and (c) correspond to 10 bit-plane coding, while (b) and (d) correspond to a 0.4 bpv target rate.*

for the MR volume MR-BRAINSAG 256x256x128. The diagrams refer to the four resulting combinations when considering two coding-rates and two slicing directions.

The selected slicing directions are the original one, which is perpendicular to the sagittal plane (128 slices along $n_2$), and the coronal one (256 slices along $n_0$). The sagittal anatomical symmetry leads to a U-shaped distortion main trend along the $n_2$ axis which is essentially linked to the space occupied by the MR brain signal with respect to the noisy background, at each slicing position. The $sP_{SNR}$ oscillation due to the 9/7 filers is clearly visible. The same oscillatory behavior appears along the $n_0$ slicing direction as shown in Fig.10(c) and (d) where few slices are considered allowing a more detailed view. Even if not shown, along $n_1$ we oserved the same things. The $sP_{SNR}$ oscillation is regular, very similar to that of Fig.8 and of the order of 0.6-0.8 dB for the 9/7 filters. On the contrary, using the 10/18 filters the oscillation contracts to less than 0.1 dB, blurring itself into the data dependent (entropic) fluctuations. At this point we also want to show that the oscillating behavior doesn't depend on particular choices of the coding rates or algorithmic coding steps. To show this, we selected a first test point corresponding to the tenth bit-plane completion by the coding algorithm [4] and a nearby one, which falls

[4] In this case we coded 10 bit-planes over a total of 14. The resulting DWT coefficient range is strongly data-dependent. Considering a rounding approximation of

in a non-specific bitstream location between the 9-th and the 10-th bit-plane. Fig.10(a) and (c) are related to the first test point, while Fig.10(b) and (d) correspond to the second one. As it will be shown in the next section, variations of the coding rate does not substantially change the nature and amplitude of the $sP_{SNR}$ fluctuation for a wide range of rates of interest.// From Fig.10 we also discover that 10/18 filters perform better than 9/7 filters not only for the oscillation amplitude but also in terms of mean value, in contrast with the high resolution predictions shown in Fig.8. As already anticipated this is due to better approximation properties of 10/18 filters with respect to the data data considered here, which in turn influences the reconstruction error statistics and allows higher performance.

### *5.2 Visual perception of the $sP_{SNR}$ oscillations*

Volumetric static (medical) data-set are usually visualized, on wide and high definition displays that reproduce a lightbox where slices appears joined to form a single view. Another usual way to see such data is the movie loop where slices are "played" with a low reproduction rate or more commonly by manual navigation. In this section we try to determine in which measure $sP_{SNR}$ oscillations are related to unpleasant visual effects when, as stated, consecutive slices are displayed one near or after another. Moreover, due to the relevance of the subject, we are also interested in what happen in the thickening case of motion compensated wavelet video coding.

Due to the intrinsic difficulty to give numerical results about the visual perception, we give observations and example that, according to our experience, can be taken as representative of different issues and situations.

- First of all, we never saw any $sP_{SNR}$ related visual effects when using even length wavelet filters (such as 10/18 or 22/14).
- When the odd length 9/7 filters are used potentially visible effects arises, which visibility depends on data, display conditions and devices and on the viewer expertise. In the case of medical data our general suggestion is to avoid using 9/7 filters. In our view, producing coded data with a significant $sP_{SNR}$ oscillation can decrease the overall quality and reliability of the data itself. For example, with respect to an agreed compression level, apparently hidden visual effects could reveal themselves in some detailed data observations, or post-processing results could be affected by the quality oscillation.
- When the short odd length filters (see Fig.8) are used the visual effects of the induced $sP_{SNR}$ oscillation are always noticeable, even at a non-expert

the wavelet coefficients to the nearest integer an 8 bit/pixel image generates a DWT coefficient range of 14 bits.

sight. We observed objectionable effects both in the case of static volumes and video coding. In the following we document those cases.

In Fig.11 we show 4 contiguous slices of the 256x256x128 MR data-set MRBRAINCOR. Slices are taken along the $n_1$ (i.e. $y$) direction. The 3D wavelet transform is calculated using the 9/3 filter bank along each direction. The coding rate is 0.27 bpv (compression ratio CR=30). Fig.12 refers to the same decoded volume but sliced along the $n_0$ (i.e. $x$) direction. Differences between upper and lower images can be quite easily perceived on display and on printed paper (with adequate printing quality). Upper images are more detailed while the lower ones are more blurred. Even if, at a first sight, this quality difference could not be considered significative, we observed that when we allow expert (physicians or image processing professionals) or non-expert viewers to look through the various slice of the whole volume with a manual back and forth movie facility, they all was able, in few seconds, to state that there was a quality alternation between consecutive slices. In Fig.13 (see the two upper signals) we show the $sP_{SNR}$ signals measured on the decoded data-set according to the two considered slicing directions. In order to give a more clear figure we restricted the viewing area to a slice interval (from slice nr. 70 to nr. 140).

We complete our analysis on visual impacts of the use of odd length wavelet filters, by considering a representative case of wavelet video coding. Wavelet based video coding intrinsically enables a series of interesting features such as spatio-temporal and quality scalability that are of particular interest for today multimedia and networking applications. Moreover efficient Motion Compensation Temporal Filtering (MCTF) solutions [21] allowed Scalable wavelet video Coding to reach performances comparable to that of the best video coding standards (AVC-H.264). Typical decoded videos generated by wavelet video codecs present periodic oscillations in the time direction (here the only one direction of interest). The bottom trace in Fig.13 reports the $sP_{SNR}$ (i.e. the frame by frame $PSNR$) in the case of a decoded test sequence. We used our wavelet based scalable video codec [32] on the Harbour sequence (in QCIF and YUV 4:2:0 format) at at the coding rate of 92kbps. The 3D decomposition is not purely dyadic but a complete temporal decomposition (MCTF with 5/3 integer lifting wavelet decomposition) is followed by a 9/7 dyadic wavelet spatial decomposition performed on each temporal subband frame. Such decomposition structure is typical in video coding systems, where the 5/3 filters are commonly used for their recognized good matching between a reliable motion estimation (very similar to the bidirectional motion estimation of hybrid schemes) and the availability of related efficient MCTF implementations [21]. The $sP_{SNR}$ oscillation of Fig.13 evidences the presence of CS phenomena of a significant entity (compatible with the use of 5/3 filters). However, because of the differences between our assumptions and the video coding architecture, the oscillation pattern differs from that issued by our 3D model. We quickly itemize such differences, but first we show how visually can differ two con-

secutive frames due to the $sP_{SNR}$ variations. This is shown Fig.14 where two original and the relating decoded frames (at 92 kpbs)are presented. The quality disparity is evident and the $PSNR$ jump in this case is of the order of 5dB indeed. The two selected frames are representative of the worse situation, however, quality alternation is actually visible along the whole video when manual frame selection or very low frame rate reproduction are used.

A quantitative analysis of the video coding case study is beyond the scope of this work. A series of factors can apparently influence the CS pattern oscillation (e.g. the decomposition structure, the MCTF, different basis for spatial and temporal filtering, different subband weighing), however the more significant one is the common use of an integer to integer non-linear lifting wavelet decomposition [5] which has been demonstrated to introduce CS error fluctuations [8]. In this case fluctuations are due to the non-linear rounding noise propagation effects produced into the lifting computation structure. Even if in [8] CS fluctuations are not described in the data domain, we can infer that when integer to integer transforms are used two ciclostationarity contributions (due to linear and non-linear effects) adds up to produce the overall CS oscillation pattern.

### *5.3 Bit-plane model degradation*

We propose a final test in order to find experimental evidence of the oscillation model degradation described in Sec.4. In particular we want to show how some features of the data to be coded have an influence on the $sP_{SNR}$ behavior. We premise that a correct interpretation of the following coding results must consider all aspects deriving from (21)-(25) at the same time. In Fig.15(a) we analyze the case of the MR-BRAINCOR volume, with a $xyz$ dimension of 256x256x64 voxels. 9/7 filers on 5 levels of 3D wavelet decomposition have been used. We considered the bit-planes $b$ from 1 to 12 and observed the $sP_{SNR}$ values after each bit-plane coding completion. The $sP_{SNR}$ signal was calculated along the MR scanning axis $n_2$ (or $z$). For the higher bit-planes 11 and 12 we can recognize the high resolution condition: the expected $sP_{SNR}$ fluctuations (see 3.2) and the 6dB distance of the curves. In the range from 5 to 10 bit-planes the $sP_{SNR}$ curves become closer owing to (24). The CS oscillation diminishes but another pseudo-periodic fluctuation gains ground, with a sort of symmetry center on the slice 34. This effect does not find an explanation in some anatomical symmetry (the considered MR-BRAINCOR data-set is coronal, i.e. acquired along the frontal direction), but it is rather

[5] The use of Integer to Integer Wavelet Transforms allows to preserve memory and decrease the computational cost. Both these are critical factors when considering the processing of a huge quantity of data, as in the case of video coding.

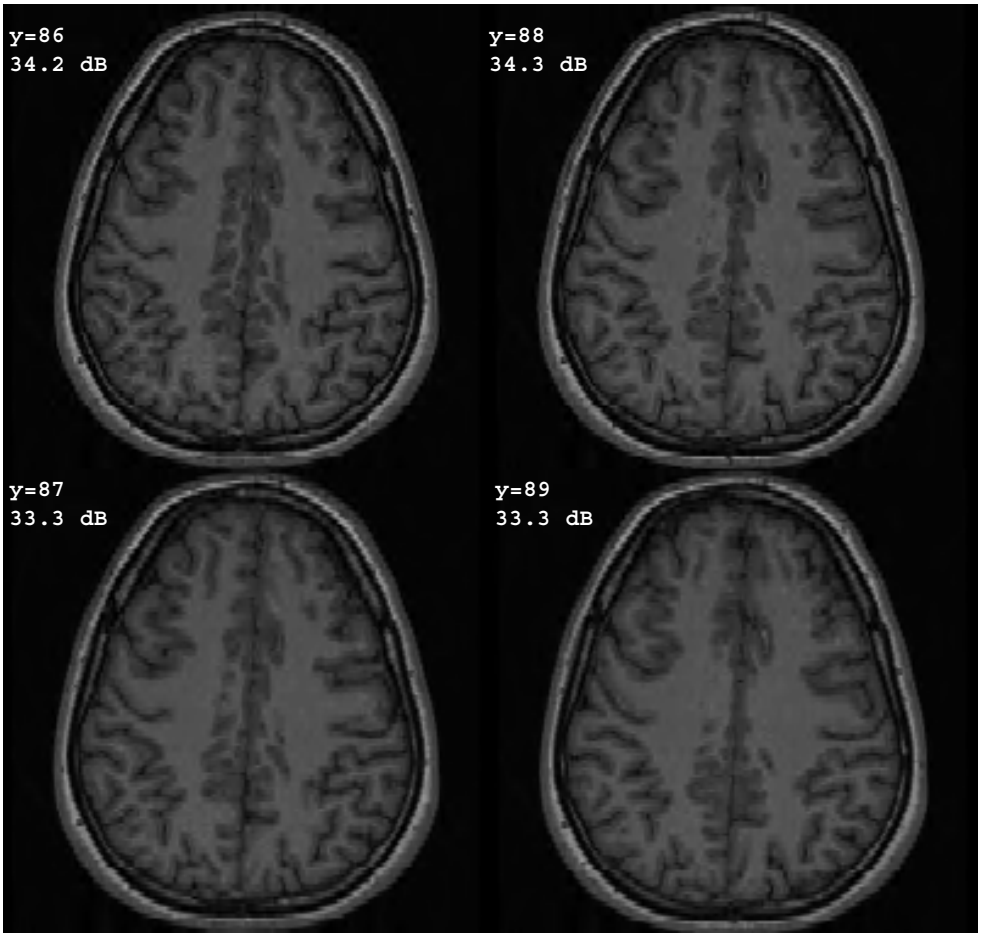


Fig. 11. *Visual results on MR-*BRAINCOR *along the y slicing direction.*

due to an interpolation related effect. In fact the 64 slices volume has been intentionally obtained from a 141 slices original one. This subsampling implies an "entropic alteration" of the slices which are governed by the resampling ratio. We don't want to study the details of this effect, but only observe that in this case, as a consequence of (25), the CS fluctuations become gradually secondary and another effect may dominate.

In Fig.15(b) we give another example in which a CT-ABDOMEN 256x256x64 volume decomposition is considered, still using 9/7 filters and a 5 level decomposition. In this case only CS related fluctuations take place because (original, non interpolated, data has been used). Obviously the short-range CS variations are superimposed to the trend determined by the sensibility of the coding algorithm to the data "information content" on every slicing plane. This long-range fluctuation trend is a passive outcome and does not affect the visual quality of salient information because the coding effort is not made intentionally selective. From Fig.15(b) we can observe that, as far as the bit-plane $sP_{SNR}$ distance is concerned, the high-resolution conditions are not satisfied from the beginning, while the amplitude of the CS fluctuations have a value

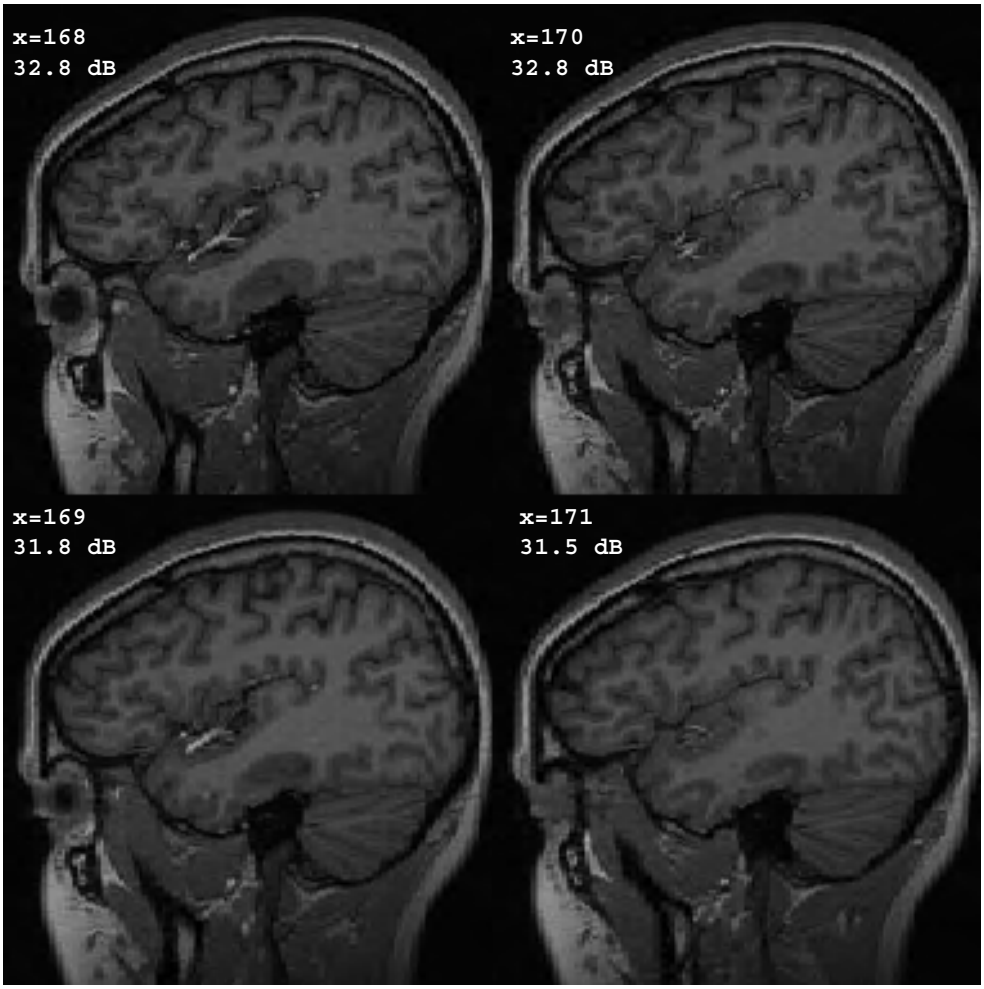


Fig. 12. *Visual results on MR-*BRAINCOR *along the x slicing direction.*

around 0.8dB (concording with the high-resolution prediction) for several $b$ values (from 9 to 12). Then the fluctuation model degrades in a different way for the two considered, equally sized, volumes. This evidences and confirm the hypothesis of a data depended model degradation, where high-frequency content or noise level are determinant features. In general and also in our case, CT data-set are less noisy compared to the MR ones. As a consequence CT related high-level detail subband histograms are more concentrated on the zero peak, and this determines a populated set of subbands for which (24) and (25) holds with inequality, even for near-lossless coding conditions. As mentioned the $sP_{SNR}$ oscillation reduces its amplitude but does not remain negligible until the eighth bit-plane is reached. In our experience the rates of interest for the lossy coding of MR and CT data-sets falls between the 9-th and the 11-th bit-plane, thus at the peak of the oscillation.

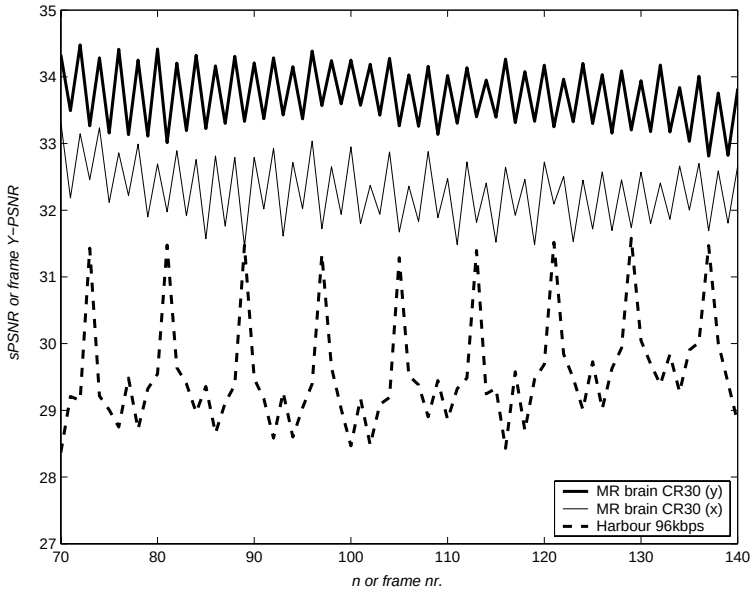


Fig. 13. *$sP_{SNR}$ oscillations on the MR-*BRAINCOR *volume and a test video.*

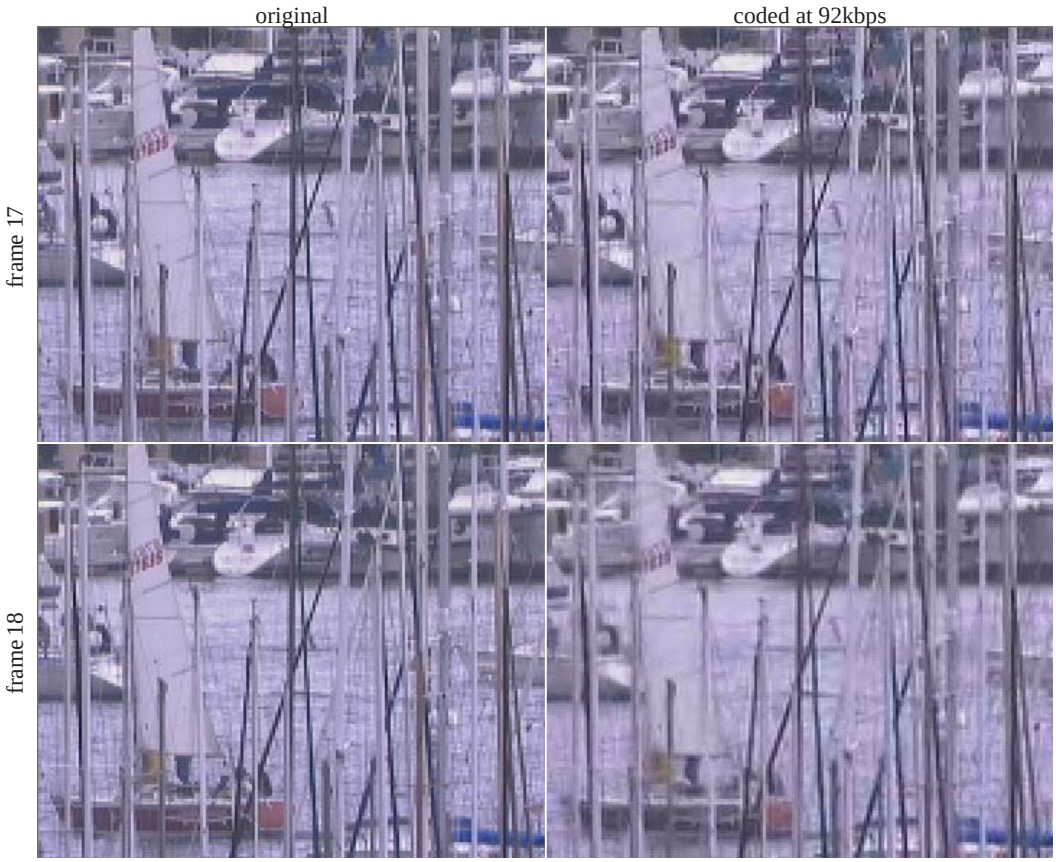


Fig. 14. *Visual results on a test video.*

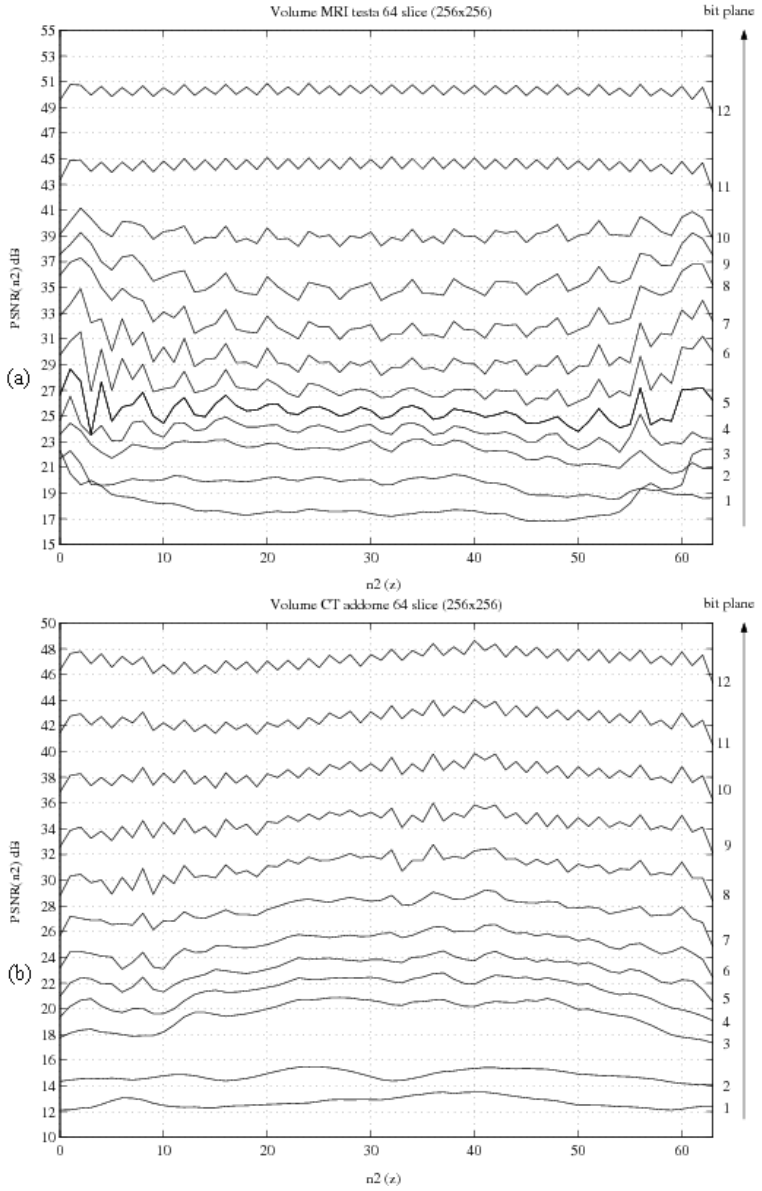


Fig. 15. *$sP_{SNR}$ curves as a function of the bit-plane number b, for 9/7 filters, along the scanning axis of the 64 slice volumes* MR-BRAINCOR *(a) and* CT-ABDOMEN *(b).*

## 6 Conclusions

Being the volume slices at adjacent position quite similar one with respect to the other, an appreciable $P_{SNR}$ oscillation due to the reconstruction process may be critical at all rates, causing objectionable artifacts and/or lowering the reliability of the coding process. In the first part of this work, the simplifying hypothesis on the quantization error was used to highlight the fundamental CS nature of the reconstruction error and to give a correct interpretation to the presence of a $P_{SNR}$ oscillation. We have observed that a structural property

of biorthogonal even length filters allows to contain such oscillation, making this filter class more attractive e.g. for medical data coding. In addition this paper has proposed a degradation analysis of the quantization error and data models towards low bit-rate conditions. It has been shown that the oscillation phenomenon is gradually reduced but is still present at medium to low bit-rates.

...

## References


[1] P. P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice Hall, Englewood Cliffs, 1993.

[2] N. Uzun, R. Haddad, Cyclostationary modeling, analysis, and optimal compensation of quantization errors in subband codecs, IEEE Trans.SP SP-43(9) (Sep.1995) 2109–2119.

[3] P. H. Westerink, J. Biemond, D. E. Boekee, Scalar quantization error analysis for image subband coding using qmf's, IEEE Trans. on Signal Proc. 40 (2) (1992) 421–428.

[4] J. Kovacevic, Subband coding systems incorporating quantizer models, IEEE Trans. on Image Proc. 4 (1995) 543–553.

[5] A. N. Delopoulos, S. D. Kollias, Optimal filter banks for signal reconstruction from noisy subband components, IEEE Trans. on Signal Proc. 44 (2) (1996) 212–224.

[6] K. Gosse, P. Duhamel, Perfect reconstruction versus MMSE filter banks in source coding, IEEE Trans. on Signal Processing 45 (9) (1997) 2188–2201.

[7] M. G. Strinzis, D. Tzovaras, Optimal construction of subband coders using Lloyd-Max quantizers, IEEE Tran. on Image Proc. 7 (5) (1998) 649–667.

[8] J. Reichel, G. Menegaz, M. Nadenau, M. Kunt, Integer wavelet transform for embedded lossy to lossless image compression, IEEE Trans.on Image Processing IP-10(3) (Mar. 2001) 383–392.

[9] P. Shelkens, A. Munteanu, An overview of volumetric coding techniques, JTC1/SC29/WG1 N2613 (2002).

[10] A. Signoroni, R. Leonardi, Diagnostic compression of biomedical volumes, in: EUSIPCO 2000, Tampere, Finland, Sep.2000, pp. 533–536.

[11] A. Signoroni, F. Lazzaroni, M. Arrigoni, R. Leonardi, Improving SPIHT-based compression of volumetric medical data, in: PCS 2001, Seoul, Korea, Apr.2001, pp. 187–190.

[12] F. Lazzaroni, A. Signoroni, R. Leonardi, Embedded morphological dilation coding for 2d and 3d images, in: VCIP 2002, Vol. SPIE(4671), San José, California, 2002, pp. 923–934.

[13] A. Bilgin, G. Zweig, M. W. Marcellin, Three-dimensional image compression using integer wavelet transforms, Appl.Opt. 39 (Apr.2000) 1799–1814.

[14] Z. Xiong, X. Wu, D. Y. Yun, W. A. Pearlman, Progressive coding of medical volumetric data using three-dimensional integer wavelet packet transforms, in: VCIP'99, Vol. SPIE-3653, Jan.1999, pp. 327–335.

[15] ISO/IEC, Information technology - jpeg2000 image coding system - part 10: Extensions for three-dimensional data and floating point data, ISO/IEC WD15444-10, working Draft (2002).

[16] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image coding using wavelet transform, IEEE Trans.IP IP-1(2) (Apr.1992) 205–220.

[17] M. J. Tsai, J. D. Villasenor, F. Chen, Stack-run image coding, IEEE Trans.CSVT CSVT-6(5) (Oct.1996) 519–521.

[18] D. Wei, H. Pai, A. Bovik, Antisymmetric biorthogonal coiflets for image coding, in: ICIP'98, Vol. 2, 1998, pp. 282–286.

[19] I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

[20] F. Lazzaroni, R. Leonardi, A. Signoroni, High-performance embedded morphological wavelet coding, in: IWDC 2002, Capri, Italy, 2002, pp. 319–326.

[21] J.-R. Ohm, Multimedia Communication Technology, Springer Verlag, 2003.

[22] N. J. Jayant, P. Noll, Digital Coding of Waveforms, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[23] M. Vetterli, J. Kovačević, Wavelets and Subband Coding, Englewood Cliffs, NJ: Prentice-Hall, 1995.

[24] S. Mallat, A Wavelet tour of Signal Processing, $2^{nd}$ Edition, Academic Press, 1999.

[25] J. M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans.SP SP-41(12) (Dec.1993) 3445–3462.

[26] A. Said, W. A. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans.CSVT CSVT-6(3) (Jun.1996) 243–250.

[27] S.-T. Hsiang, J. W. Woods, Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling, in: MPEG-4 Workshop and Exhibition at ISCAS 2000, Geneva, Switzerland, 2000.

[28] B. B. Chai, J. Vass, X. Zhuang, Significance-linked connected component analysis for wavelet image coding, IEEE Trans. Image Processing 8 (1999) 774–784.

[29] D. Taubman, High performance scalable image compression with EBCOT, Trans. Image Processing 9 (2000) 1158–1170.

[30] S. Mallat, F. Falzon, Analysis of low bit rate image transform coding, IEEE Trans.SP SP-46(4) (Apr.1998) 1027–1042.

[31] C. E. Pearson (Ed.), Handbook of applied mathematics, selected results and methods, $2^{nd}$ Edition, VNR, New York, 1990.

[32] N. Adami, M. Brescianini, M. Dalai, R. Leonardi, A. Signoroni, A fully scalable video coder with inter-scale wavelet prediction and morphological coding, in: VCIP 2005, Beijing, Cina, 2005, accepted contribution.