# PERFECT RECONSTRUCTION IN REDUCED REDUNDANCY WAVELET-BASED MULTIPLE DESCRIPTION CODING OF IMAGES

*Teodora Petrişor*[1]*, Béatrice Pesquet-Popescu*[1] *and Jean-Christophe Pesquet*[2]

(1) ENST, Signal and Image Processing Department, 46, rue Barrault, 75634 Paris Cédex 13, France
(2) IGM and UMR-CNRS 8049, Université de Marne-la-Vallée,
5, Bd Descartes, Champs sur Marne, 77454 Marne la Vallée Cédex 2, France.
Email: {petrisor,pesquet}@tsi.enst.fr,pesquet@univ-mlv.fr

## ABSTRACT

In this paper, we consider frame expansions derived from biorthogonal wavelet bases for building multiple descriptions with low redundancy constraints. Such constraints rise the problem of perfect reconstruction of the associated decompositions in the absence of quantization or channel errors, which requires special attention and therefore will be detailed in this work. We will show that several schemes that yield perfect reconstruction are possible with the proposed strategy. Moreover, when the resulting coefficients are corrupted by quantization or channel errors, we employ a fast iterative algorithm based on projections onto convex sets in order to enhance the quality of the decoded images.

## 1. INTRODUCTION

Building multiple descriptions for image transmission over error-prone channels without priority mechanisms, such as Internet and cellular networks, has proved to be a very promising error resilience technique [1]. A common problem in these networks is that of the transient channel shutdowns, due to either network congestion, accompanied by packet losses, or to deep error fades in wireless networks. The idea behind multiple description coding is to create two or several correlated representations of the source and take advantage of the often available path diversity in communication systems, in order to provide the user with a minimum reconstruction quality attainable when only one description is received and enhanceable when several descriptions are combined. The descriptions can either be built in the raw domain of the data (voice, image, video) or from some transformed versions of it, the latter approach being the most common [2], [3], [4].

In this work we address the problem of building descriptions using a wavelet *frame decomposition* of a two-dimensional signal. In particular, biorthogonal 9-7 wavelets are employed in a corresponding oversampled filter bank structure in order to create two descriptions of an image. This structure has an inherent redundancy which can be exploited into descriptions, in the same time providing us with shift invariance [5]. The proposed wavelet frame has an additional feature which is the reduced redundancy achieved through a further quincunx subsampling of the detail subbands. An important problem when reducing the redundancy is that the resulting representations may no longer provide perfect recovery [6], [7], [8]. We shall address this issue by considering the polyphase transfer matrices for the proposed schemes and studying their invertibility in the frequency domain.

The outline of the paper is as follows: Section 2 presents the framework for building two wavelet descriptions, highlighting the fact that in the two-dimensional case several possible schemes can be built for the proposed wavelet frame. The problem of perfect reconstruction is addressed in Section 3 by a frequency-domain study of the polyphase transfer matrix. Section 4 presents an optimization technique based on iterative projections for central decoder design. In Section 5 we provide some simulation results.

## 2. MULTIPLE DESCRIPTION FRAMEWORK

The proposed frame decomposition is derived from a classical biorthogonal wavelet basis. Consider the impulse responses $(h[n])_{n\in\mathbb{Z}}$ and $(g[n])_{n\in\mathbb{Z}}$ of the analysis low-pass and high-pass filter banks respectively, associated to such a decomposition. When applied to each of the dimensions, we obtain the following subband coefficients:

$$\begin{aligned} a[n,m] &= \sum_{k,l} x[k,l]h[2n-k]h[2m-l] \\ dh[n,m] &= \sum_{k,l} x[k,l]h[2n-k]g[2m-l] \\ dv[n,m] &= \sum_{k,l} x[k,l]g[2n-k]h[2m-l] \\ dd[n,m] &= \sum_{k,l} x[k,l]g[2n-k]g[2m-l] \end{aligned} \tag{1}$$

where $a$ stands for the approximation subband coefficients at a given resolution, $dh$, $dv$, $dd$ denote the horizontal, vertical and diagonal detail subband coefficients at the same resolution level and $x$ is the approximation sequence at the previous finer resolution. For simplicity, we have omitted here the resolution level index $j$ which varies between 1 and $J \in \mathbb{N}^*$.

Note that the two-dimensional filter bank applied in (1) is dyadic and separable. We propose an overcomplete representation by considering a second decomposition of the form:

$$\begin{aligned} a_{(s,s')}[n,m] &= \sum_{k,l} x[k,l]h[2n+s-k]h[2m+s'-l] \\ dh_{(s,s')}[n,m] &= \sum_{k,l} x[k,l]h[2n+s-k]g[2m+s'-l] \\ dv_{(s,s')}[n,m] &= \sum_{k,l} x[k,l]g[2n+s-k]h[2m+s'-l] \\ dd_{(s,s')}[n,m] &= \sum_{k,l} x[k,l]g[2n+s-k]g[2m+s'-l] \end{aligned} \tag{2}$$

where we have introduced the shift parameters $s$ and $s'$, which belong to $\{0,1\}$. It should be noted that, in our approach, for $j < J$ we keep the usual non redundant decomposition whereas, for the last resolution level, Eq. (2) is used

where $s$ and $s'$ are not equal to 0 simultaneously. In other words, the redundancy is only introduced at the last level of the subband decomposition.

Transmitting the entire set of coefficients resulting from Eqs. (1) and (2) would still introduce a high redundancy. This would obviously lead to a high reconstruction quality in any situation (both descriptions being received or only one of them) but such a solution remains computationally and ratewise virtually unacceptable. Therefore we propose a subsampled version of this structure, having only a slightly higher redundancy than the critically subsampled decomposition.

In the subsampled schemes that we build, not all choices of $s$ and $s'$ lead to frame decompositions, but we shall prove in the following that some of them do. Having a frame decomposition enables the perfect reconstruction of the signal in the absence of quantization or transmission errors (channel noise).

Let us now propose several multiple description schemes each corresponding to a specific subsampling in Eqs. (1) and (2). Our main guideline is to build two balanced descriptions, each containing some of the coefficients from each representation (the original one and the shifted one).

In order to provide acceptable side reconstructions, we aim at keeping the main image features in each description. Therefore we shall not perform a further subsampling on the approximation subbands in each of the representations. In turn, the detail subbands are quincunx subsampled. In this manner the overall redundancy only depends on the last level of decomposition and amounts to the number of additional approximation coefficients. Thus, the coarser the last resolution level is, the lower the redundancy we get.

Recall that the quincunx sampling of a 2-D field $(x[n,m])_{n,m}$ leads to two polyphase components:

$$x^{(q)}[n,m] = x[n+m+q, n-m] \tag{3}$$

where $q \in \{0,1\}$.

As noticed in the former section, the two representations (1), (2) are identical except for the last level of decomposition. Considering Eq. (3), we build two descriptions as follows:

1. **Description I** is formed by the set of coefficients $C_J^{\mathrm{I}}$ and the detail subbands $\{dh_j^{(0)}, dv_j^{(0)}, dd_j^{(0)}\}$ defined at resolution levels $j \in \{1,\ldots,J-1\}$;
2. **Description II** contains the set of coefficients $C_J^{\mathrm{II}}$ as well the other quincunx polyphase components of each detail subband: $\{dh_j^{(1)}, dv_j^{(1)}, dd_j^{(1)}\}$, for $j \in \{1,\ldots,J-1\}$.

Here, we have denoted by $C_J^{\mathrm{I}}$ (resp. $C_J^{\mathrm{II}}$) the set of all subband coefficients at the coarsest resolution in the first (resp. second) description. These sets will be of the form:

$$C_J^{\mathrm{I}} = \{a_{J,(0,0)}, dh^{(p_1)}_{J,(r_1,r'_1)}, dv^{(p_2)}_{J,(r_2,r'_2)}, dd^{(p_3)}_{J,(r_3,r'_3)}\}$$
$$C_J^{\mathrm{II}} = \{a_{J,(s,s')}, dh^{(p_4)}_{J,(r_4,r'_4)}, dv^{(p_5)}_{J,(r_5,r'_5)}, dd^{(p_6)}_{J,(r_6,r'_6)}\}$$

where $p_i \in \{0,1\}$, $i \in \{1,\ldots,6\}$, denotes the selected quincunx polyphase component for the $i$-th detail coefficient sequence at resolution level $J$. Also, for all $i \in \{1,\ldots,6\}$, we have either $(r_i,r'_i) = (0,0)$ or $(r_i,r'_i) = (s,s')$. In other words, at the coarsest level, the detail sequences in each description are obtained as quincunx subsampled detail subbands from the original and one shifted decomposition.

In this paper we consider two of the possible overcomplete expansions, based on translated filters. The first one is given by $s = s' = 1$, and we shall denote it later on by the index $(1,1)$. In this case we obtained only two possible sets of wavelet subbands that provide perfect reconstruction. These are given by the whole critically sampled decomposition from Eq. (1) to which we added the approximation subband from Eq. (2) or the similar structure considering all of the second basis coefficients and the approximation from the first basis as redundancy. These schemes are obviously completely recoverable in the absence of quantization since they include the critically sampled decomposition. Moreover, the study that we conducted led to the conclusion that the reconstruction error in these two redundant schemes is below the one obtained from the critically sampled decomposition. In this manner the central decoder does exploit the introduced redundancy in order to increase the quality of the reconstruction. It is worth noting that the so-obtained combinations do not facilitate the building of balanced descriptions, as will be shown in Fig. 3.

A more interesting case is when $s = 1 - s'$. For each of these combinations we obtain at least 12 schemes that can be perfectly recoverable. These 12 schemes also yield a smaller reconstruction error as compared with the critically sampled scheme. In the next section we present the framework that led to these conclusions.

## 3. PERFECT RECONSTRUCTION

### 3.1 Polyphase formulation

By discarding some of the detail coefficients, the global system no longer has a frame structure for all combinations of polyphase components in the detail subbands. It is therefore important to identify the combinations which ensure perfect reconstruction. To this end we study the polyphase transfer matrix of our system.

By passing into the frequency domain and using matrix notations, we re-write the convolutions/decimations in Eqs. (1) and (2). Let

$$\boldsymbol{M}_0(\ \ ) = \begin{bmatrix} H_0(\ \ ) & H_1(\ \ ) \\ G_0(\ \ ) & G_1(\ \ ) \end{bmatrix} \tag{4}$$

be the polyphase matrix corresponding to the filter bank operating along one of the dimensions. $H_0$ and $H_1$ are the two polyphase components of $H$:

$$\begin{aligned} H_0(\ \ ) &= \frac{1}{2}[H(\tfrac{}{2}) + H(\tfrac{}{2} + \ \ )] \\ H_1(\ \ ) &= \frac{e^{\frac{\imath}{2}}}{2}[H(\tfrac{}{2}) - H(\tfrac{}{2} + \ \ )] \end{aligned} \tag{5}$$

and similar notations are used for $G$. Shifting by 1 the impulse responses leads to a polyphase transfer matrix of the form:

$$\boldsymbol{M}_1(\ \ ) = \begin{bmatrix} H_1(\ \ ) & e^{\imath} & H_0(\ \ ) \\ G_1(\ \ ) & e^{\imath} & G_0(\ \ ) \end{bmatrix}. \tag{6}$$

Thus, the polyphase transfer matrix for the 2D separable representation in Eq. (1) or Eq. (2) is given by the Kronecker tensor product: $\mathscr{M}_{(r,r')}(\ \ x,\ \ y) = \boldsymbol{M}_r(\ \ x) \otimes \boldsymbol{M}_{r'}(\ \ y)$, where

$(r,r') = (0,0)$ or $(r,r') = (s,s')$. The global transfer equation of our system can be written as:

$$\begin{bmatrix} \mathscr{C}_{(0,0)}(\ x,\ y) \\ \mathscr{C}_{(s,s')}(\ x,\ y) \end{bmatrix} = \begin{bmatrix} \mathscr{M}_{(0,0)}(\ x,\ y) \\ \mathscr{M}_{(s,s')}(\ x,\ y) \end{bmatrix} \mathscr{X}(\ x,\ y) \qquad (7)$$

where $\mathscr{X}(\ x,\ y)$ is the vector of the Fourier transforms of the 4 polyphase components of the input image: $(x[2n-k,2m-l])_{n,m}$ with $(k,l) \in \{0,1\}^2$. The subband coefficient vector on the left-hand side of Eq. (7) contains the 4-dimensional vector of the Fourier transforms of the coefficients of the first representation in the upper part and from the second one in the lower part. By putting emphasis on the quincunx polyphase components of the coefficients as defined by Eq. (3), Eq. (7) can be rewritten under the form:

$$\begin{bmatrix} \mathscr{C}^{(0)}_{(0,0)}(\ x,\ y) \\ \mathscr{C}^{(1)}_{(0,0)}(\ x,\ y) \\ \mathscr{C}^{(0)}_{(s,s')}(\ x,\ y) \\ \mathscr{C}^{(1)}_{(s,s')}(\ x,\ y) \end{bmatrix} = \begin{bmatrix} \widetilde{\mathscr{M}}_{(0,0)}(\ x,\ y) \\ \widetilde{\mathscr{M}}_{(s,s')}(\ x,\ y) \end{bmatrix} \begin{bmatrix} \mathscr{X}^{(0)}(\ x,\ y) \\ \mathscr{X}^{(1)}(\ x,\ y) \end{bmatrix} \qquad (8)$$

where, for $(r,r') \in \{(0,0),(s,s')\}$,

$$\widetilde{\mathscr{M}}_{(r,r')}(\ x,\ y) = \begin{bmatrix} \widetilde{\mathscr{M}}^{(0)}_{(r,r')}(\ x,\ y) & \widetilde{\mathscr{M}}^{(1)}_{(r,r')}(\ x,\ y) \\ \widetilde{\mathscr{M}}^{(1)}_{(r,r')}(\ x,\ y) & e^{\imath(\ x+\ y)}\widetilde{\mathscr{M}}^{(0)}_{(r,r')}(\ x,\ y) \end{bmatrix} \qquad (9)$$

with

$$\widetilde{\mathscr{M}}^{(0)}_{(r,r')}(\ x,\ y) = \frac{1}{2}(\mathscr{M}_{(r,r')}(\ x,\ y) + \mathscr{M}_{(r,r')}(\ x+\ ,\ y+\ ))$$
$$\widetilde{\mathscr{M}}^{(1)}_{(r,r')}(\ x,\ y) = \frac{e^{\imath\ x}}{2}(\mathscr{M}_{(r,r')}(\ x,\ y) - \mathscr{M}_{(r,r')}(\ x+\ ,\ y+\ ))$$

with $\ x = (\ x+\ y)/2$ and $\ y = (\ x-\ y)/2$. In the left-hand side of Eq. (8), we end up with a subband coefficient vector having 16 components while both vectors $\mathscr{X}^{(q)}(\ x,\ y)$, $q \in \{0,1\}$, have 4 components. Among the subband coefficient components, we shall keep only 10: the 4 approximation components and 6 detail ones. Once this choice has been made, let us denote by $\bar{\mathscr{M}}(\ x,\ y)$ the submatrix of size $10 \times 8$ formed by the corresponding selected lines of the polyphase transfer matrix in Eq. (8). The perfect reconstruction of the proposed scheme is guaranteed if and only if $\bar{\mathscr{M}}(\ x,\ y)$ is left-invertible for all $(\ x,\ y) \in [0,2\ )^2$. In the following we designate this matrix as the quincunx polyphase transfer matrix.

### 3.2 Invertibility of the polyphase transfer matrix

The left invertibility of the polyphase transfer matrix can be studied by considering its singular value decomposition. A necessary and sufficient condition for perfect reconstruction is that none of its eight singular values vanishes on the unit bi-circle. For the three possible combinations of $s$ and $s'$ in Eq. (2), we have studied the evolution on the unit bi-circle of the minimum singular values of each quincunx polyphase transfer matrix corresponding to one of the considered low-redundancy schemes. It must be pointed out that all the results presented subsequently have been obtained by using classical 9-7 biorthogonal filters.

Considering the shift of the filter impulse responses by (1,1), we show in Fig. 1 the variation w.r.t. frequency of the minimum singular value of the matrix $\bar{\mathscr{M}}(\ x,\ y)$. The invertibility of the system is guaranteed, since for all $(\ x,\ y) \in [0,2\ )$ the minimum singular value is nonzero. In Fig. 2 we also show a less obvious combination of polyphase quincunx detail subbands, that yields perfect reconstruction, as well as a combination that does not.

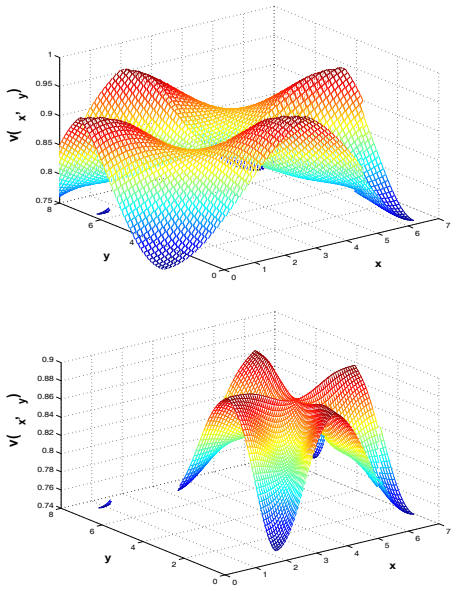


Figure 1: Minimum singular value $v$ of the quincunx polyphase transfer matrix as a function of frequency for schemes: $D^{\mathrm{I}} \cup D^{\mathrm{II}} = \{a_{(0,0)}, dh_{(0,0)}, dv_{(0,0)}, dd_{(0,0)}, a_{(1,1)}\}$ (top) and the critically sampled decomposition (bottom).

For two of the schemes that provide us with perfect reconstruction, we build the two descriptions, as explained in Section 2. These will be denoted by $D^{\mathrm{I}}$ and $D^{\mathrm{II}}$ in Figs. 1 and 2. The first scheme corresponds to $(s,s') = (1,1)$ and it has the following distribution of the coefficients between the two descriptions at the last level: $D^{\mathrm{I}}_{(1,1)} = \{a_{(0,0)}, dh^{(0)}_{(0,0)}, dv^{(0)}_{(0,0)}, dd^{(0)}_{(0,0)}\}$ and $D^{\mathrm{II}}_{(1,1)} = \{a_{(1,1)}, dh^{(1)}_{(0,0)}, dv^{(1)}_{(0,0)}, dd^{(1)}_{(0,0)}\}$. In this case, the perfect reconstruction that is reflected by Fig. 1 can be deduced more directly by observing that the $a_{(1,1)}$ approximation sequence comes in addition to the decomposition onto a basis and thus the overall decomposition is clearly invertible.

The second perfect reconstruction scheme is obtained with $(s,s') = (0,1)$ and it is formed by the following descriptions: $D^{\mathrm{I}}_{(0,1)} = \{a_{(0,0)}, dh^{(0)}_{(0,1)}, dv^{(0)}_{(0,0)}, dd^{(0)}_{(0,0)}\}$ and $D^{\mathrm{II}}_{(0,1)} = \{a_{0,1}, dh^{(1)}_{0,1}, dv^{(1)}_{0,0}, dd^{(1)}_{0,0}\}$. This combination also leads to a smaller reconstruction error than in the critically sampled case.

It can be noticed that the study of the singular values of the matrix $\bar{\mathscr{M}}(\ x,\ y)$ in each situation provides a means to evaluate the mean square reconstruction error. By assuming that the quantization noise is white, it can be shown that the two proposed oversampled schemes lead to a smaller reconstruction error than a critically subsampled 9-7 filter bank structure.

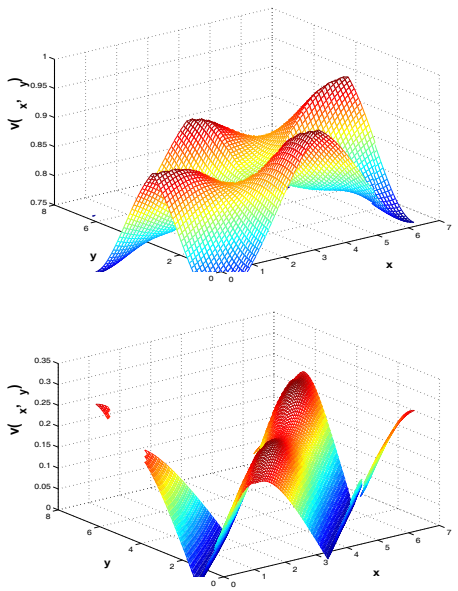


Figure 2: Minimum singular value $v$ of the quincunx polyphase transfer matrix as a function of frequency for schemes: $D^{\mathrm{I}} \cup D^{\mathrm{II}} = \{a_{(0,0)}, dh_{(0,1)}, dv_{(0,0)}, dd_{(0,0)}, a_{(0,1)}\}$ (top) and of one of the combinations that do not yield perfect reconstruction (bottom).

## 4. OPTIMIZED LOSSY RECONSTRUCTION

At the central decoder we employ a fast iterative algorithm that enhances the quality of the reconstruction. This algorithm is useful both when the two descriptions are received or when only one of them is available. This algorithm detailed in [9], is based on the minimization of a convex quadratic objective function under convex constraints. The main idea is to consider the quantization constraints as prior information on the image, corresponding to convex sets:

$$S_j = \{x \mid -\frac{}{2} \leq f[j]^{\mathsf{T}} x - \hat{} [j] \leq \frac{}{2}\} \qquad (10)$$

where is the quantization step, $x$ is the original image, $\hat{} [j]$ is a quantized wavelet coefficient and $f[j]$ is a vector corresponding to the function used to compute this coefficient. The optimal solution will be a point in the intersection of all the so-defined convex sets. It is given by the projection on this intersection of a reference image $x_0$ which is an initial estimate of the original one. In the case of the central decoder we choose it as a weighted sum of the recovered images after decoding the two descriptions.

The iterative algorithm allows for parallel computing and offers fast convergence properties.

## 5. EXPERIMENTAL RESULTS

We present some of the test results for the schemes $D_{(1,1)}$ and $D_{(0,1)}$, on the $512 \times 512$ gray scale "Barbara" image. We have performed a three-level wavelet decomposition with biorthogonal 9-7 filter banks for both schemes. The quantized coefficients have been encoded with the EZBC coder, [10]. In Fig. 3 we have plotted the rate-distortion curves for the two schemes before the iterative reconstruction (*Init. Scheme*) and after it (*Opt. Scheme*). In the top graph we present the central decoders for Scheme 1 ($D_{(1,1)}$) and

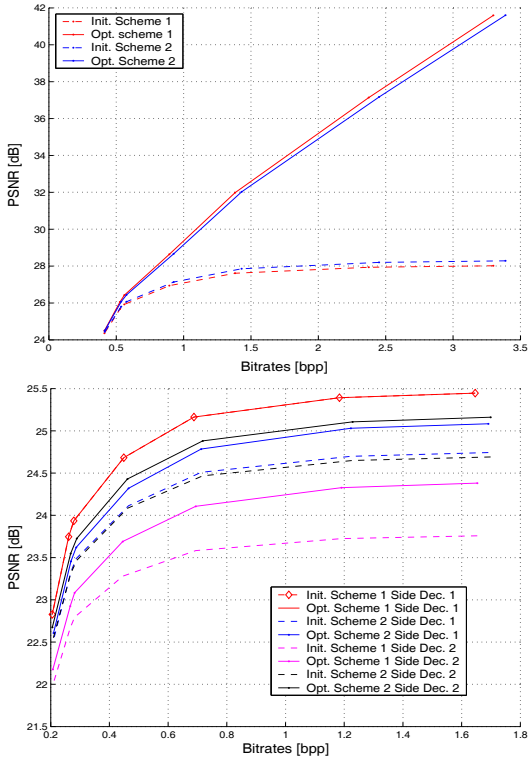


Figure 3: Rate-Distortion evaluation of the two schemes for "Barbara": central (top graph) and side (bottom graph) decoders.

Scheme 2 ($D_{(0,1)}$). They exhibit similar performances. In the second graph is highlighted the importance of having balanced descriptions. The two side decoders of Scheme 2 exhibit lower performances than the first side decoder of Scheme 1, but closer quality between each other, which makes for a preferable tradeoff.

### REFERENCES

[1] V.K. Goyal, J. Kovačević, R. Arean, and M. Vetterli, "Multiple description transform coding of images," in *IEEE Int. Conf. on Image Processing*, Chicago, IL, 1998, pp. 674–678.

[2] D.-M. Chung and Y. Wang, "Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 895–908, September 1999.

[3] I. V. Bajic and J. W. Woods, "Domain-based multiple description coding of images and video," *IEEE Transactions on Image Processing*, vol. 12(18), pp. 1211–1225, October 2003.

[4] R. Motwani and C. Guillemot, "Tree-structured oversampled filterbanks as joint source-channel codes: application to image transmission over erasure channels," *IEEE Transactions on Signal Processing*, vol. 52 (9), pp. 2584 – 2599, Sept. 2004.

[5] J.-C. Pesquet, H. Krim, and H. Carfantan, "Time invariant orthonormal wavelet representations," *IEEE Transactions on Signal Processing*, vol. 40(10), pp. 1964–1970, 1996.

[6] H. Bolcskei, F. Hlawatsch, and H.G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Transactions on Signal Processing*, vol. 46(12), pp. 3256–3268, December 1998.

[7] T.D. Tran, R. L. de Quieroz, and T. Q. Nguyen, "Linear-phase perfect reconstruction filter bank: lattice structure, design and application in image coding," *IEEE Transactions on Signal Processing*, vol. 48(1), pp. 133–147, January 2000.

[8] T. Tanaka and Y. Yamashita, "On perfect reconstruction with lost channel data in lapped pseudo-orthogonal transform," in *EUSIPCO*, 2004, pp. 877–880.

[9] T. Petrisor, B. Pesquet-Popescu, and J.-C. Pesquet, "Wavelet-based multiple description coding of images with iterative convex optimization techniques," to appear in *IEEE Int. Conf. on Image Processing*, 2005.

[10] S. Hsiang and J. Woods, "Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling," in *Int. Symp. Circ. and Syst.*, Geneva, May 2000, pp. 589–593.