Research Article

Block-Based Adaptive Vector Lifting Schemes for Multichannel Image Coding

Amel Benazza-Benyahia,[1] Jean-Christophe Pesquet,[2] Jamel Hattay,[1] and Hela Masmoudi[3, 4]

[1] *Unité de Recherche en Imagerie Satellitaire et ses Applications (URISA), Ecole Supérieure des Communications (SUP'COM), Tunis 2083, Tunisia*

[2] *Institut Gaspard Monge and CNRS-UMR 8049, Université de Marne la Vallée, 77454 Marne la Vallée Cédex 2, France*

[3] *Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA*

[4] *US Food and Drug Administration, Center of Devices and Radiological Health, Division of Imaging and Applied Mathematics, Rockville, MD 20852, USA*

We are interested in lossless and progressive coding of multispectral images. To this respect, nonseparable vector lifting schemes are used in order to exploit simultaneously the spatial and the interchannel similarities. The involved operators are adapted to the image contents thanks to block-based procedures grounded on an entropy optimization criterion. A vector encoding technique derived from EZW allows us to further improve the efficiency of the proposed approach. Simulation tests performed on remote sensing images show that a significant gain in terms of bit rate is achieved by the resulting adaptive coding method with respect to the non-adaptive one.

1. INTRODUCTION

The interest in multispectral imaging has been increasing in many fields such as agriculture and environmental sciences. In this context, each earth portion is observed by several sensors operating at different wavelengths. By gathering all the spectral responses of the scene, a multicomponent image is obtained. The spectral information is valuable for many applications. For instance, it allows pixel identification of materials in geology and the classification of vegetation type in agriculture. In addition, the long-term storage of such images is highly desirable in many applications. However, it constitutes a real bottleneck in managing multispectral image databases. For instance, in the Landsat 7 Enhanced Thematic Mapper Plus system, the 8-band multispectral scanning radiometer generates 3.8 Gbits per scene with a data rate of 150 Mbps. Similarly, the Earth Orbiter I (EO-I) instrument works at a data bit rate of 500 Mbps. The amount of data will continue to become larger with the increase of the number of spectral bands, the enhancement of the spatial resolution, and the improvement of the radiometry accuracy requiring finer quantization steps. It is expected that the next

Landsat generation will work at a data rate of several Gbps. Hence, compression becomes mandatory when dealing with multichannel images. Several methods for data reduction are available, the choice strongly depend on the underlying application requirements [1]. Generally, on-board compression techniques are lossy because the acquisition data rates exceed the downlink capacities. However, ground coding methods are often lossless so as to avoid distortions that could damage the estimated values of the physical parameters corresponding to the sensed area. Besides, scalability during the browsing procedure constitutes a crucial feature for ground information systems. Indeed, a coarse version of the image is firstly sent to the user to make a decision about whether to abort the decoding if the data are considered of little interest or to continue the decoding process and refine the visual quality by sending additional information. The challenge for such progressive decoding procedure is to design a compact multiresolution representation. Lifting schemes (LS) have proved to be efficient tools for this purpose [2, 3]. Generally, the 2D LS is handled in a separable way. Recent works have however introduced nonseparable quincunx lifting schemes (QLS) [4]. The QLS can be viewed as the next

generation of coders following nonrectangularly subsampled filterbanks [5–7]. These schemes are motivated by the emergence of quincunx sampling image acquisition and display devices such as in the SPOT5 satellite system [8]. Besides, nonseparable decompositions offer the advantage of a "true" two-dimensional processing of the images presenting more degrees of freedom than the separable ones. A key issue of such multiresolution decompositions (both LS and QLS) is the design of the involved decomposition operators. Indeed, the performance can be improved when the intrinsic spatial properties of the input image are accounted for. A possible adaptation approach consists in designing space-varying filter banks based on conventional adaptive linear mean square algorithms [9–11]. Another solution is to adaptively choose the operators thanks to a nonlinear decision rule using the local gradient information [12–15]. In a similar way, Taubman proposed to adapt the vertical operators for reducing the edge artifacts especially encountered in compound documents [16]. Boulgouris et al. have computed the optimal predictors of an LS in the case of specific wide-sense stationary fields by considering an a priori autocovariance model of the input image [17]. More recently, adaptive QLS have been built without requiring any prior statistical model [8] and, in [18], a 2D orientation estimator has been used to generate an edge adaptive predictor for the LS. However, all the reported works about adaptive LS or QLS have only considered monocomponent images. In the case of multicomponent images, it is often implicitly suggested to decompose separately each component. Obviously, an approach that takes into account the spectral similarities in addition to the spatial ones should be more efficient than the componentwise approach. A possible solution as proposed in Part 2 of the JPEG2000 standard [19] is to apply a reversible transform operating on the multiple components before their spatial multiresolution decomposition. In our previous work, we have introduced the concept of *vector* lifting schemes (VLS) that decompose *simultaneously* all the spectral components in a separable manner [20] or in a nonseparable way (QVLS) [21]. In this paper, we consider blockwise adaptation procedures departing from the aforementioned adaptive approaches. Indeed, most of the existing works propose a pointwise adaptation of the operators, which may be costly in terms of bit rate.

More precisely, we propose to firstly segment the image into nonoverlapping blocks which are further classified into several regions corresponding to different statistical features. The QVLS operators are then optimally computed for each region. The originality of our approach relies on the optimization of a criterion that operates directly on the entropy, which can be viewed as a sparsity measure for the multiresolution representation.

This paper is organized as follows. In Section 2, we provide preliminaries about QVLS. The issue of the adaptation of the QVLS operators is addressed in Section 3. The objective of this section is to design efficient adaptive multiresolution decompositions by modifying the basic structure of the QVLS. The choice of an appropriate encoding technique is also discussed in this part. In Section 4, experimental results are presented showing the good performance of the
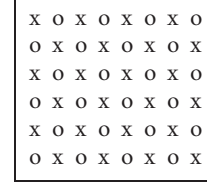

Figure 1: Quincunx sampling grid: the polyphase components $x_0^{(b)}(m, n)$ correspond to the "x" pixels whereas the polyphase components $\tilde{x}_0^{(b)}(m, n)$ correspond to the "o" pixels.

proposed approach. A comparison of the fixed and variable block size strategies is also performed. Finally, some concluding remarks are given in Section 5.

2. VECTOR QUINCUNX LIFTING SCHEMES

2.1. The lifting principle

In a generic LS, the input image is firstly split into two sets $\mathcal{S}_1$ and $\mathcal{S}_2$ of spatial samples. Because of the local correlation, a predictor (P) allows to predict the $\mathcal{S}_1$ samples from the $\mathcal{S}_2$ ones and to replace them by their prediction errors. Finally, the $\mathcal{S}_2$ samples are smoothed using the residual coefficients thanks to an update (U) operator. The updated coefficients correspond to a coarse version of the input signal and, a multiresolution representation is then obtained by recursively repeating this decomposition to the updated approximation coefficients. The main advantage of the LS is its reversibility regardless of the choice of the P and U operators. Indeed, the inverse transform is simply obtained by reversing the order of the operators (U-P) and substituting a minus (resp., plus) sign by a plus (resp., minus) one. Thus, the LS can be considered as an appealing tool for exact and progressive coding. Generally, the LS is applied to images in a separable manner as for instance in the 5/3 wavelet transform retained for the JPEG2000 standard.

2.2. Quincunx lifting scheme

More general LS can be obtained with nonseparable decompositions giving rise to the so-called QLS [4]. In this case, the $\mathcal{S}_1$ and $\mathcal{S}_2$ sets, respectively, correspond to the two quincunx polyphase components $x_{j/2}^{(b)}(m, n)$ and $\tilde{x}_{j/2}^{(b)}(m, n)$ of the approximation $a_{j/2}^{(b)}(m, n)$ of the bth band at resolution $j/2$ (with $j \in \mathbb{N}$):

$$
\begin{aligned}
x_{j/2}^{(b)}(m, n) &= a_{j/2}^{(b)}(m - n, m + n), \\
\tilde{x}_{j/2}^{(b)}(m, n) &= a_{j/2}^{(b)}(m - n + 1, m + n),
\end{aligned}
\tag{1}
$$

where (m, n) denotes the current pixel. The initialization is performed at resolution $j = 0$ by taking the polyphase components of the original image $x(n, m)$ when this one has been rectangularly sampled (see Figure 1). We have then $a_0(n, m) = x(n, m)$. If the quincunx subsampled version of the original image is available (e.g., in the SPOT5 system), the initialization of the decomposition process is performed at
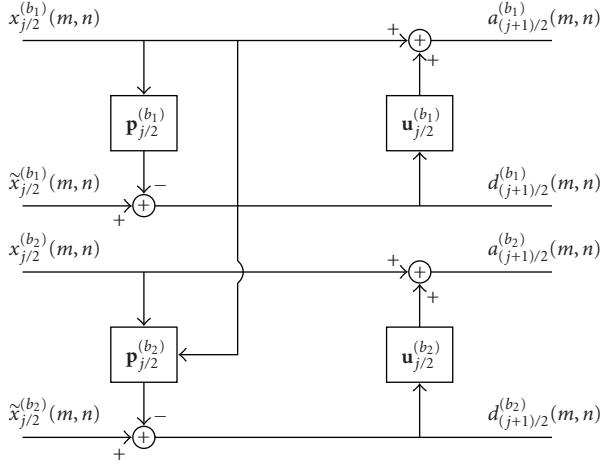
FIGURE 2: An example of a decomposition vector lifting scheme in the case of a two-channel image.

resolution $j = 1/2$ by setting $a_{1/2}^{(b)}(n,m) = x^{(b)}(m-n, m+n)$. In the P step, the prediction errors $d_{(j+1)/2}^{(b)}(m,n)$ are computed:

$$d_{(j+1)/2}^{(b)}(m,n) = \tilde{x}_{j/2}^{(b)}(m,n) - \lfloor \mathbf{x}_{j/2}^{(b)}(m,n)^\top \mathbf{p}_{j/2}^{(b)} \rceil, \qquad (2)$$

where $\lfloor \cdot \rceil$ is a rounding operator, $\mathbf{x}_{j/2}^{(b)}(m,n)$ is a vector containing some $a_{j/2}^{(b)}(m,n)$ samples, and, $\mathbf{p}_{j/2}^{(b)}$ is a vector of prediction weights of the same size. The approximation $a_{(j+1)/2}^{(b)}(m,n)$ of $a_{j/2}^{(b)}(m,n)$ is an updated version of $x_{j/2}^{(b)}(m,n)$ using some of the $d_{(j+1)/2}^{(b)}(m,n)$ samples regrouped into the vector $\mathbf{d}_{j/2}^{(b)}(m,n)$:

$$a_{(j+1)/2}^{(b)}(m,n) = x_{j/2}^{(b)}(m,n) + \lfloor \mathbf{d}_{j/2}^{(b)}(m,n)^\top \mathbf{u}_{j/2}^{(b)} \rceil, \qquad (3)$$

where $\mathbf{u}_{j/2}^{(b)}$ is the associated update weight vector. The resulting approximation can be further decomposed so as to get a multiresolution representation of the initial image. Unlike classical separable multiresolution analyses where the input signal is decimated by a factor 4 to generate the approximation signal, the number of pixels is divided by 2 at each (half-) resolution level of the nonseparable quincunx analysis.

2.3. Vector quincunx lifting scheme

The QLS can be extended to a QVLS in order to exploit the interchannel redundancies in addition to the spatial ones. More precisely, the $d_{j/2}^{(b)}(m,n)$ and $a_{j/2}^{(b)}(m,n)$ coefficients are now obtained by using coefficients of the considered band b *and* also coefficients of the *other* channels. Obviously, the QVLS represents a versatile framework, the QLS being a special case. Besides, the QVLS is quite flexible in terms of selection of the prediction mask and component ordering. Figure 2 shows the corresponding analysis structures. As an example of particular interest, we will consider the simple QVLS whose P operator relies on the following neighbors of

the coefficient $a_{j/2}^{(b)}(m-n+1, m+n)$:

$$\mathbf{x}_{j/2}^{(b_1)}(m,n) = \begin{pmatrix} a_{j/2}^{(b_1)}(m-n, m+n) \\ a_{j/2}^{(b_1)}(m-n+1, m+n-1) \\ a_{j/2}^{(b_1)}(m-n+1, m+n+1) \\ a_{j/2}^{(b_1)}(m-n+2, m+n) \end{pmatrix},$$

$$\forall i > 1, \quad \mathbf{x}_{j/2}^{(b_i)}(m,n) = \begin{pmatrix} a_{j/2}^{(b_i)}(m-n, m+n) \\ a_{j/2}^{(b_i)}(m-n+1, m+n-1) \\ a_{j/2}^{(b_i)}(m-n+1, m+n+1) \\ a_{j/2}^{(b_i)}(m-n+2, m+n) \\ \hline a_{j/2}^{(b_{i-1})}(m-n+1, m+n) \\ \vdots \\ a_{j/2}^{(b_1)}(m-n+1, m+n) \end{pmatrix},$$

$$(4)$$

where $(b_1, \ldots, b_B)$ is a given permutation of the channel indices $(1, \ldots, B)$. Thus, the component b_1, which is chosen as a reference channel, is coded by making use of a purely spatial predictor. Then, the remaining components b_i (for $i > 1$) are predicted both from neighboring samples of the same component b_i (spatial mode) *and* from the samples of the previous components b_k (for $k < i$) located at the same position. The final step corresponds to the following update, which is similarly performed for all the channels:

$$\mathbf{d}_{j/2}^{(b_i)}(m,n) = \begin{pmatrix} d_{j/2}^{(b_i)}(m-1, n+1) \\ d_{j/2}^{(b_i)}(m,n) \\ d_{j/2}^{(b_i)}(m-1, n) \\ d_{j/2}^{(b_i)}(m, n+1) \end{pmatrix}. \qquad (5)$$

Note that such a decomposition structure requires to set $4B + (B-1)B/2$ parameters for the prediction weights and $4B$ parameters for the update weights. It is worth mentioning that the update filter feeds the cross-channel information back to the approximation coefficients since the detail coefficients contain information from other channels. This may appear as an undesirable situation that may lead to some leakage effects. However, due to the strong correlation between the channels, the detail coefficients of the B channels have a similar frequency content and no quality degradation was observed in practice.

3. ADAPTATION PROCEDURES

3.1. Entropy criterion

The compression ability of a QVLS-based representation depends on the appropriate choice of the P and U operators. In general, the mean entropy $\mathcal{H}_J$ is a suitable measure of compactness of the J-stage multiresolution representation. This measure which is independent of the choice of the encoding

algorithm is defined as the average of the entropies $\mathcal{H}_J^{(b)}$ of the B channel data:

$$\mathcal{H}_J \triangleq \frac{1}{B} \sum_{b=1}^{B} \mathcal{H}_J^{(b)}. \tag{6}$$

Likewise, $\mathcal{H}_J^{(b)}$ is calculated as a weighted average of the entropies of the approximation and the detail subbands:

$$\mathcal{H}_J^{(b)} \triangleq \left(\sum_{j=1}^{J} 2^{-j} \mathcal{H}_{d,j/2}^{(b)} \right) + 2^{-J} \mathcal{H}_{a,J/2}^{(b)}, \tag{7}$$

where $\mathcal{H}_{d,j/2}^{(b)}$ (resp., $\mathcal{H}_{a,J/2}^{(b)}$) denotes the entropy of the detail (resp., approximation) coefficients of the bth channel, at resolution level $j/2$.

3.2. Optimization criteria

As mentioned in Section 1, the main contribution of this paper is the introduction of some adaptivity rules in the QVLS schemes. More precisely, the parameter vectors $\mathbf{p}_{j/2}^{(b)}$ are modified according to the local activity of each subband. For this purpose, we have envisaged block-based approaches which start by partitioning each subband of each spectral component into blocks. Then, for a given channel b, appropriate classification procedures are applied in order to cluster the blocks which can use the same P and U operators within a given class $c \in \{1, \dots, C_{j/2}^{(b)}\}$. It is worth pointing out that the partition is very flexible as it depends on the considered spectral channel. In other words, the block segmentation yields different maps from a channel to another. In this context, the entropy $\mathcal{H}_{d,j/2}^{(b)}$ is expressed as follows:

$$\mathcal{H}_{d,j/2}^{(b)} = \sum_{c=1}^{C_{j/2}^{(b)}} \pi_{j/2}^{(b,c)} \mathcal{H}_{d,j/2}^{(b,c)}, \tag{8}$$

where $\mathcal{H}_{d,j/2}^{(b,c)}$ denotes the entropy of the detail coefficients of the bth channel within class c and, the weighting factor $\pi_{j/2}^{(b,c)}$ corresponds to the probability that a detail sample $d_{j/2}^{(b)}$ falls into class c. Two problems are subsequently addressed: (i) the optimization of the QVLS operators, (ii) the choice of the block segmentation method.

3.3. Optimization of the predictors

We now explain how a specific statistical modeling of the detail coefficients within a class c can be exploited to efficiently optimize the prediction weights. Indeed, the detail coefficients $d_{(j+1)/2}^{(b)}$ are often viewed as realizations of a continuous zero mean random variable X whose probability density function f is given by a generalized Gaussian distribution

(GGD) [22, 23]:

$$\forall x \in \mathbb{R}, \quad \mathsf{f}\left(x; \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)$$
$$= \frac{\beta_{(j+1)/2}^{(b,c)}}{2\alpha_{(j+1)/2}^{(b,c)} \Gamma\left(1/\beta_{(j+1)/2}^{(b,c)}\right)} e^{-\left(|x|/\alpha_{(j+1)/2}^{(b,c)}\right)^{\beta_{(j+1)/2}^{(b,c)}}}, \tag{9}$$

where $\Gamma(z) \triangleq \int_0^{+\infty} t^{z-1} e^{-t} dt$, $\alpha_{(j+1)/2}^{(b,c)} > 0$ is the scale parameter, and $\beta_{(j+1)/2}^{(b,c)} > 0$ is the shape parameter. These parameters can be easily estimated from the empirical moments of the data samples [24]. The GGD model allows to express the differential entropy $\mathcal{H}(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)})$ as follows:

$$\mathcal{H}\left(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)$$
$$= \log\left(\frac{2\alpha_{(j+1)/2}^{(b,c)} \Gamma\left(1/\beta_{(j+1)/2}^{(b,c)}\right)}{\beta_{(j+1)/2}^{(b,c)}} \right) + \frac{1}{\beta_{(j+1)/2}^{(b,c)}}. \tag{10}$$

It is worth noting that the proposed lifting structure generates integer-valued coefficients that can be viewed as quantized versions of the continuous random variable X with a quantization step $q = 1$. According to high rate quantization theory [25], the differential entropy $\mathcal{H}(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)})$ provides a good estimate of $\mathcal{H}_{d,j/2}^{(b,c)}$. In practice, the following empirical estimator of the detail coefficients entropy is employed:

$$\widehat{\mathcal{H}}_{d,K_{j/2}^{(b,c)}}\left(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)$$
$$= -\frac{1}{K_{j/2}^{(b,c)}} \sum_{k=1}^{K_{j/2}^{(b,c)}} \log\left(\mathsf{f}\left(\widetilde{x}_{j/2}^{(b,c)}(k) - \left(\mathbf{x}_{j/2}^{(b,c)}(k)\right)^{\top} \mathbf{p}_{j/2}^{(b,c)}\right) \right), \tag{11}$$

where $\widetilde{x}_{j/2}^{(b,c)}(1), \dots, \widetilde{x}_{j/2}^{(b,c)}(K_{j/2}^{(b,c)})$ and $\mathbf{x}_{j/2}^{(b,c)}(1), \dots, \mathbf{x}_{j/2}^{(b,c)}(K_{j/2}^{(b,c)})$ are $K_{j/2}^{(b,c)} \in \mathbb{N}^*$ realizations of $\widetilde{x}_{j/2}^{(b)}$ and $\mathbf{x}_{j/2}^{(b)}$ classified in c. As we aim at designing the most compact representation, the objective is to compute the predictor $\mathbf{p}_{j/2}^{(b,c)}$ that minimizes $\mathcal{H}_J$. From (6), (7), and (8), it can be deduced that the optimal parameter vector also minimizes $\mathcal{H}_{d,j/2}^{(b)}$ and therefore, $\mathcal{H}(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)})$, which is consistently estimated by $\widehat{\mathcal{H}}_{d,K_{(j+1)/2}^{(b,c)}}(\alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)})$. This leads to the maximization of

$$L\left(\mathbf{p}_{j/2}^{(b,c)}; \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)$$
$$= \sum_{k=1}^{K_{j/2}^{(b,c)}} \log\left(\mathsf{f}\left(\widetilde{x}_{j/2}^{(b,c)}(k) - \left(\mathbf{x}_{j/2}^{(b,c)}(k)\right)^{\top} \mathbf{p}_{j/2}^{(b,c)}\right) \right). \tag{12}$$

Thus, the maximum likelihood estimator of $\mathbf{p}_{j/2}^{(b,c)}$ must be determined. From (9), we deduce that the optimal predictor minimizes the following $\ell^{\beta_{(j+1)/2}^{(b,c)}}$ criterion:

$$\ell^{\beta_{(j+1)/2}^{(b,c)}}\left(\mathbf{p}_{j/2}^{(b,c)}; \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)$$
$$\triangleq \sum_{k=1}^{K_{j/2}^{(b,c)}} \left| \widetilde{x}_{j/2}^{(b,c)}(k) - \left(\mathbf{x}_{j/2}(k)^{(b,c)}\right)^{\top} \mathbf{p}_{j/2}^{(b,c)} \right|^{\beta_{(j+1)/2}^{(b,c)}}. \tag{13}$$

Hence, thanks to the GGD model, it is possible to design a predictor in each class c that ensures the compactness of the representation in terms of the resulting detail subband entropy. However, it has been observed that the considered statistical model is not always adequate for the approximation subbands which makes impossible to derive a closed form expression for the approximation subband entropy. Related to this fact, several alternatives can be envisaged for the selection of the update operator. For instance, it can be adapted to the contents of the image so as to minimize the reconstruction error [8]. It is worth noticing that, in this case, the underlying criterion is the variance of the reconstruction error and not the entropy. A simpler alternative that we have retained in our experiments consists in choosing the same update operator for all the channels, resolution levels, and clusters. Indeed, in our experiments, it has been observed that the decrease of the entropy is mainly due to the optimization of the predictor operators.

3.4. Fixed-size block segmentation

The second ingredient of our adaptive approach is the block segmentation procedure. We have envisaged two alternatives. The first one consists in iteratively classifying *fixed* size blocks as follows [8].

INIT

The block size $s_{j/2}^{(b)} \times t_{j/2}^{(b)}$ and the number of regions $C_{j/2}^{(b)}$ are fixed by the user. Then, the approximation $a_{j/2}^{(b)}$ is partitioned into nonoverlapping blocks that are classified into $C_{j/2}^{(b)}$ regions. It should be pointed out that the classification of the approximation subband has been preferred to that of the detail subbands at a given resolution level j. Indeed, it is expected that homogenous regions (in the spatial domain) share a common predictor, and such homogeneous regions are more easily detected from the approximation subbands than from the detail ones. For instance, a possible classification map can be obtained by clustering the blocks according to their mean values.

PREDICT

In each class c, the GGD parameters $\alpha_{(j+1)/2}^{(b,c)}$ and, $\beta_{(j+1)/2}^{(b,c)}$ are estimated as described in [24]. Then, the optimal predictor $\mathbf{p}_{j/2}^{(b,c)}$ that minimizes the $\ell^{\beta_{(j+1)/2}^{(b,c)}}$ criterion is derived. The initial values of the predictor weights are set by minimizing the detail coefficient variance.

ASSIGN

The contents of each class c are modified so that a block of details initially in class c could be moved to another class c^* according to some assignment criterion. More precisely, the global entropy $\mathcal{H}_{d,j/2}^{(b,c)}$ is equal to the sum of the contributions of all the detail blocks within class c. This additive property enables to easily derive the optimal assignment rule. At each

resolution level and, according to the retained band ordering, a current block $\mathbf{B}$ is assigned to a class c^* if its contribution to the entropy of that class induces the maximum decrease of the global entropy. This amounts to move the block $\mathbf{B}$, initially assumed to belong to class c, to class c^* if the following condition is satisfied:

$$h\left(\mathbf{B}, \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right) < h\left(\mathbf{B}, \alpha_{(j+1)/2}^{(b,c^*)}, \beta_{(j+1)/2}^{(b,c^*)}\right), \qquad (14)$$

where

$$
\begin{aligned}
&h\left(\mathbf{B}, \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right) \\
&\triangleq \sum_{m=1}^{s_{j/2}^{(b)}} \sum_{n=1}^{t_{j/2}^{(b)}} \log\left(\mathsf{f}\left(\mathbf{B}(m,n); \alpha_{(j+1)/2}^{(b,c)}, \beta_{(j+1)/2}^{(b,c)}\right)\right).
\end{aligned}
\qquad (15)
$$

PREDICT and *ASSIGN* steps are repeated until the convergence of the global entropy. Then, the procedure is iterated through the J resolution stages.

At the convergence of the procedure, at each resolution level, the chosen predictor for each block is identified with a binary index code which is sent to the decoder leading to an overall overhead not exceeding

$$o = \left\lceil \sum_{b=1}^{B} \sum_{j=1}^{J} \frac{\log_2\left(C_{j/2}^{(b)}\right)}{s_{j/2}^{(b)} t_{j/2}^{(b)}} \right\rceil \text{(bpp)}. \qquad (16)$$

Note that the amount of side information can be further reduced by differential encoding.

3.5. Variable-size block segmentation

More flexibility can be achieved by varying the block sizes according to the local activity of the image. To this respect, a quadtree (QT) segmentation in the spatial domain is used which provides a layered representation of the regions in the image. For simplicity, this approach has been implemented using a volumetric segmentation (same segmentation for each image channel at a given resolution as depicted in Figure 3) [26]. The regions are obtained according to a segmentation criterion $\mathcal{R}$ that is suitable for compression purposes. Generally, the QT can be built following two alternatives: a splitting or a merging approach. The first one starts from a partition of the transformed multicomponent image into volumetric quadrants. Then, each quadrant f is split into 4 volumetric subblocks $c_1, \ldots, c_4$ if the criterion $\mathcal{R}$ holds, otherwise the untouched quadrant f is associated with a leaf of the unbalanced QT. The subdivision is eventually repeated on the subblocks $c_1, \ldots, c_4$ until the subblock minimum size $k_1 \times k_2$ is achieved. Finally, the resulting block-shaped regions correspond to the leaves of the unbalanced QT.

In contrast, the initial step of the dual approach (i.e., the merging procedure) corresponds to a partition of the image into minimum size $k_1 \times k_2$ subblocks. Then, the homogeneity with respect to the rule $\mathcal{R}$ of each quadrant formed by adjacent volumetric subblocks $c_1, \ldots, c_4$ is checked. In case of homogeneity, the fusion of $c_1, \ldots, c_4$ is carried out, giving rise to a father block f. Similar to the splitting approach,
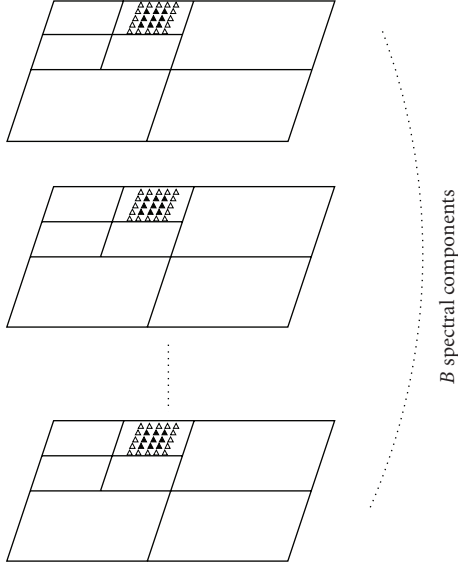
FIGURE 3: An example of a volumetric block-partitioning of a B-component image.

the fusion procedure is recursively performed until the whole image size is reached.

Obviously, the key issue of such QT partitioning lies in the definition of the segmentation rule $\mathcal{R}$. In our work, this rule is based on the lifting optimization criterion. Indeed, in the case of the splitting alternative, the objective is to decide whether the splitting of a node f into its 4 children $c_1, \ldots, c_4$ provides a more compact representation than the node f does. For each channel, the optimal prediction and update weights $\mathbf{p}_{j/2}^{(b,f)}$ $\mathbf{u}_{j/2}^{(b,f)}$ of node f are computed for a J-stage decomposition. The optimal weights $\mathbf{p}_{j/2}^{(b,c_i)}$ and, $\mathbf{u}_{j/2}^{(b,c_i)}$ of the children $c_1, \ldots, c_4$ are also computed. Let $\mathcal{H}_{d,j/2}^{(b,f)}$ and, $\mathcal{H}_{d,j/2}^{(b,c_i)}$ denote the entropy of the resulting multiresolution representations. The splitting is decided if the following inequality $\mathcal{R}$ holds:

$$\frac{1}{4B} \sum_{i=1}^{4} \left(\sum_{b=1}^{B} \mathcal{H}_{d,j/2}^{(b,c_i)} \right) + o(c_i) < \frac{1}{B} \left(\sum_{b=1}^{B} \mathcal{H}_{d,j/2}^{(b,f)} \right) + o(f), \tag{17}$$

where $o(n)$ is the coding cost of the side information required by the decoding procedure at node n. This overhead information concerns the tree structure and the operators weights. Generally, it is easy to code the QT by assigning the bit "1" to an intermediate node and the bit "0" to a leaf. Since the image corresponds to all the leaves of the QT, the problem amounts to the coding of the binary sequences pointing on these terminating nodes. To this respect, a run-length coder is used. Concerning the operators weights, these ones should be exactly coded. As they take floating values, they are rounded prior to the arithmetic coding stage. Obviously, to avoid any mismatch, the approximation and detail coefficients are computed according to these rounded weights. Finally, it is worth noting that the merging rule is derived in a straightforward way from (17).

TABLE 1: Description of the test images.

Name	Number of components	Source	Scene
Trento6	6	Thematic Mapper	Rural
Trento7	7	Thematic Mapper	Rural
Tunis3	3	SPOT3	Urban
Kair4	4	SPOT4	Rural
Tunis4-160	4	SPOT4	Rural
Tunis4-166	4	SPOT4	Rural

TABLE 2: Influence of the prediction optimization criterion on the average entropies for non adaptive 4-level QLS and QVLS decompositions. The update was fixed for all resolution levels and for all the components.

Image	QLS ℓ^2	QLS ℓ^β	Gain	QVLS ℓ^2	QVLS ℓ^β	Gain
Trento6	4.2084	4.1172	0.0912	3.8774	3.7991	0.0783
Trento7	3.9811	3.8944	0.0867	3.3641	3.2988	0.0653
Tunis3	5.3281	5.2513	0.0768	4.5685	4.4771	0.0914
Kair4	4.3077	4.1966	0.1111	3.9222	3.8005	0.1217
Tunis4-160	4.7949	4.7143	0.0806	4.2448	4.1944	0.0504
Tunis4-166	3.9726	3.9075	0.0651	3.7408	3.6205	0.1203
Average	4.4321	4.3469	0.0853	3.9530	3.8651	0.0879

3.6. Improved EZW

Once the QVLS coefficients have been obtained, they are encoded by an embedded coder so as to meet the scalability requirement. Several scalable coders exist which can be used for this purpose, for example, the embedded zerotree wavelet coder (EZW) [27], the set partitioning in hierarchical tree (SPIHT) coder [28], the embedded block coder with optimal truncation (EBCOT) [29]. Nevertheless, the efficiency of such coders can be increased in the case of multispectral image coding as will be shown next. To illustrate this fact, we will focus on the EZW coder which has the simplest structure. Note however that the other existing algorithms can be extended in a similar way.

The EZW algorithm allows a scalable reconstruction in quality by taking into account the interscale similarities between the detail coefficients [27]. Several experiments have indeed indicated that if a detail coefficient at a coarse scale is insignificant, then all the coefficients in the same orientation and in the same spatial location at finer scales are likely to be insignificant too. Therefore, spatial orientation trees whose nodes are detail coefficients can be easily built, the scanning order starts from the coarsest resolution level. The EZW coder consists in detecting and encoding these insignificant coefficients through a specific data structure called a zerotree. This tree contains elements whose values are smaller than the current threshold T_i. The use of the EZW coder results in dramatic bit savings by assigning to a zerotree a

TABLE 3: Average entropies for several lifting-based decompositions. Two resolution levels were used for the separable decompositions and four (half-)resolution levels for the nonseparable ones. The update was fixed except for Gouze's decomposition OQLS (6,4).

Image	5/3	RKLT+5/3	QLS (4,2)	OQLS (6,4)	Our QLS	Our QVLS	Merging QLS $k_1 = 16$ $k_2 = 16$	RKLT and merging QLS $k_1 = 16$ $k_2 = 16$	Merging QVLS $k_1 = 16$ $k_2 = 16$
Trento6	3.9926	3.9260	4.6034	3.9466	4.1172	3.7991	3.7243	3.5322	**3.4822**
Trento7	3.7299	3.7384	4.4309	3.9771	3.8944	3.2988	3.5543	3.3219	**3.0554**
Tunis3	5.0404	4.6586	5.7741	4.7718	5.2513	4.4771	4.2038	3.9425	**3.0998**
Kair4	4.0581	3.9104	4.6879	3.8572	4.1966	3.8005	3.6999	3.5240	**3.1755**
Tunis4-160	4.5203	4.2713	5.2312	4.1879	4.7143	4.1944	4.1208	3.6211	**3.2988**
Tunis4-166	3.6833	3.5784	4.4807	3.6788	3.9075	3.6205	3.8544	3.2198	**3.0221**
Average	4.1708	4.0138	4.8680	4.0699	4.3469	3.8651	3.8596	3.5269	**3.1890**

single symbol (ZTR) at the position of its root. In his pioneering paper, Shapiro has considered only *separable* wavelet transforms. In [30], we have extended the EZW to the case of *nonseparable* QLS by defining a modified parent-child relationship. Indeed, each coefficient in a detail subimage at level $(j + 1)/2$ is the father of *two* colocated coefficients in the detail subimage at level $j/2$. It is worth noticing that a tree rooted in the coarsest approximation subband will have one main subtree rooted in the coarsest detail subband. As in the separable case, the Quincunx EZW (QEZW) alternates between dominant passes DP_i and subordinate passes SP_i at each round i. All the wavelet coefficients are initially put in a list called the dominant list, DL_1, while the other list SL_1 (the subordinate list) is empty. An initial threshold T_1 is chosen and the first round of passes R_1 starts ($i = 1$). The dominant pass DP_i detects the significant coefficients with respect to the current threshold T_i. The signs of the significant coefficients are coded with either POS or NEG symbols. Then, the significant coefficients are set to zero in DL_i to facilitate the formation of zerotrees in the next rounds. Their magnitudes are put in the subordinate list, SL_i. In contrast, the descendants of insignificant coefficient are tested for being included in a zerotree. If this cannot be achieved, then these coefficients are isolated zeros and they are coded with the specific symbol IZ. Once all the elements in DL_i have been processed, the DP_i ends and the SP_i starts: each significant coefficient in SL_i will have a reconstruction value given by the decoder. By default, an insignificant coefficient will have a reconstruction value equal to zero. During SP_i, the uncertainty interval is halved. The new reconstruction value is the center of this smaller uncertainty range depending on whether its magnitude lies in the upper (UPP) or lower (LOW) half. Once the SL_i has been fully processed, the next iteration starts by incrementing i.

Therefore, for each channel, both EZW and QEZW provide a set of coefficients $(d_n^{(b)})_n$ encoded according to the selected scanning path. We subsequently propose to modify the QEZW algorithm so as to jointly encode the components of the B-uplet $(d_n^{(1)}, \dots, d_n^{(B)})_n$. The resulting algorithm will be designated as V-QEZW. We begin with the observation that,
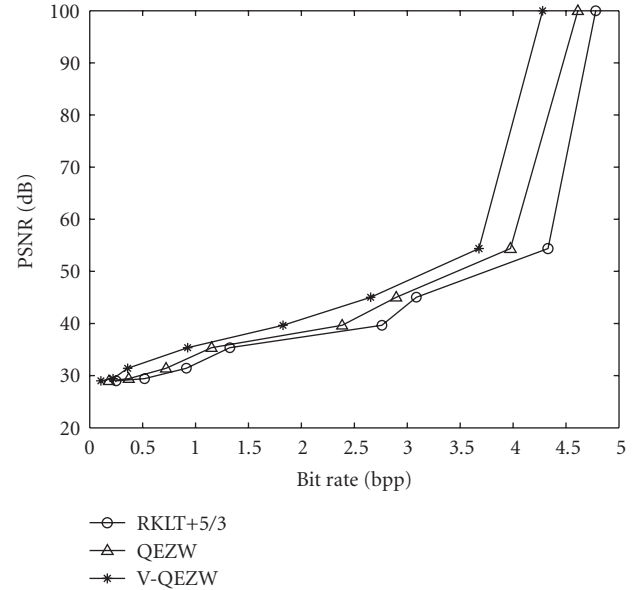

FIGURE 4: Image Trento7: average PSNR (in dB) *versus* average bit rate (in bpp) generated by the embedded coders with the equivalent number of decomposition stages. The EZW coder is associated with the RKLT+5/3 transform and the QEZW, and the V-QEZW with the same QVLS. We have adopted that the convention PSNR = 100 dB amounts to an infinite PSNR.

if a coefficient $d_n^{(b)}$ is significant with respect to a fixed threshold, then all the coefficients $d_n^{(b')}$ in the other channel $b' \neq b$ are likely to be significant with respect to the same threshold. Insignificant or isolated zero coefficients also satisfy such inter channel similarity rule. The proposed coding algorithm will avoid to manage and encode *separately* B dominant lists and B subordinate lists. The vector coding technique introduces 4 extra-symbols that indicate that for a given index n, all the B coefficients are either positive significant (APOS) or negative significant (ANEG), or insignificant (AZTR) or isolated zeros (AIZ). More precisely, at each iteration of the V-QEZW, the significance map of the b_1 channel conveys both
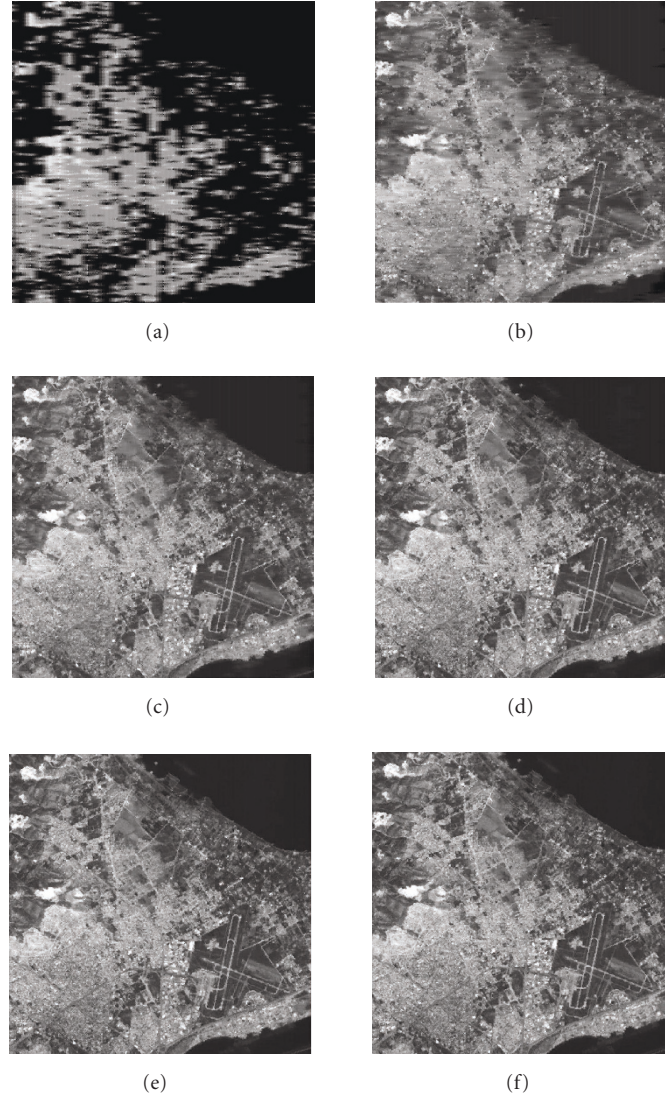
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 5: Recontructed images at several passes of the V-QEZW concerning the first channel ($b = 1$) of the SPOT image TUNIS. (a) PSNR = 21.0285 dB channel bit rate = 0.1692 bpp. (b) PSNR = 28.2918 dB channel bit rate = 0.7500 bpp. (c) PSNR = 32.9983 dB channel bit rate = 1.4946 bpp. (d) PSNR = 39.5670 dB channel bit rate = 2.4972 bpp. (e) PSNR = 57.6139 dB channel bit rate = 4.2644 bpp. (f) PSNR = $+\infty$ channel bit rate = 4.5981 bpp

inter- and intrachannel information using the 3- bit codes: APOS, ANEG, AIZ, AZTR, POS, NEG, IZ, ZTR. The remaining channel significance maps are only concerned with intrachannel information consisting of POS, NEG, IZ, ZTR symbols coded with 2 bits. The stronger the similarities are, the more efficient the proposed technique is.

4.   EXPERIMENTAL RESULTS

Table 1 lists the 512×512 multichannel images used in our experiments. All these images are 8 bpp multispectral satellite images. The Trento6 image corresponds to the Landsat-Thematic Mapper Trento7 image where the sixth component has been discarded since it is not similar to the

other components. As the entropy decrease is not significant when more than 4 (half-)resolution levels are considered, we choose to use 4-stage nonseparable decompositions ($J = 4$). All the proposed decompositions make use of a fixed update $\mathbf{u}_{j/2}^{(b)} = (1/8, 1/8, 1/8, 1/8)^{\top}$. The employed vector lifting schemes implicitly correspond to the band ordering that ensures the most compact representation. More precisely, an exhaustive search was performed for the SPOT images ($B \leq 4$) by examining all the permutations. If a greater number of components are involved as for the Thematic Mapper images, this approach becomes computationally intractable. Hence, an efficient algorithm must be applied for computing a feasible band ordering. Since more than one band are used for prediction, it is not straightforward to view the problem

as a graph theoretic problem [31]. Therefore, heuristic solutions should be found for band ordering. In our case, we have considered the correlations between the components and used the component(s) that is least correlated in an intracoding mode and the others in intercoding mode. Alternatively, the band with the smallest entropy is coded in intramode as a reference band, the others in intermode.

First of all, we validate the use of the GGD model for the detail coefficients. Table 2 gives the global entropies obtained with the QLS and the QVLS first using global minimum variance predictors, then using global GGD-derived predictors (i.e., minimizing the ℓ^β criterion in (13)). It shows that using the predictors derived from the ℓ^β criterion yields improved performance in the monoclass case. It is important to observe that, even in the nonadaptive case (one single class), the GGD model is more suitable to derive optimized predictors. Besides, Table 2 shows the outperformance of QVLS over QLS, always in the nonadaptive case. For instance, in the case of Tunis4-160, a gain of 0.52 bpp is achieved by the QVLS schemes over the componentwise QLS.

In Table 3, the variable block size adaptive versions of the proposed QLS and QVLS are compared to those obtained with the most competitive reversible wavelet-based methods. All of the latter methods are applied separately to each spectral component. In particular, we have tested the 5/3 biorthogonal transform. Besides, prior the 5/3 transform or our QLS, a reversible Karhunen-Loève transform (RKLT) [32] has been applied to decorrelate the B components as recommended in Part 2 of the JPEG2000 standard. As a benchmark, we have also retained the OQLS (6,4) reported in [8] which uses an optimized update and a minimum variance predictor. It can be noted that the merging procedure was shown to outperform the splitting one and that it leads to substantial gains for both the QLS and QVLS. Our simulations also confirm the superiority of the QVLS over the optimal spectral decorrelation by the RKLT. Figure 4 provides the variations of the average PSNR versus the average bit rate achieved at each step of the QEZW or V-QEZW coder for the Trento7 data. As expected, the V-QEZW algorithm leads to a lower bit rate than the QEZW. At the final reconstruction pass, the V-QEZW bit rate is 0.33 bpp below the QEZW one. Figure 5 displays the reconstructed images for the first channel of the Tunis3 scene, which are obtained at the different steps of the V-QEZW algorithm. These results demonstrate clearly the scalability in accuracy of this algorithm, which is suitable for telebrowsing applications.

5. CONCLUSION

In this paper we have suggested several tracks for improving the performance of lossless compression for multichannel images. In order to take advantage of the correlations between the channels, we have made use of vector-lifting schemes combined with a joint encoding technique derived from EZW. In addition, a variable-size block segmentation approach has been adopted for adapting the coefficients of the predictors of the considered VQLS structure to the local contents of the multichannel images. The gains obtained

on satellite multispectral images show a significant improvement compared with existing wavelet-based techniques. We think that the proposed method could also be useful in other imaging application domains where multiple sensors are used, for example, medical imaging or astronomy.

Note

Part of this work has been presented in [26, 33, 34].

REFERENCES

[1] K. Sayood, *Introduction to Data Compression*, Academic Press, San Diego, Calif, USA, 1996.

[2] W. Sweldens, "Lifting scheme: a new philosophy in biorthogonal wavelet constructions," in *Wavelet Applications in Signal and Image Processing III*, vol. 2569 of *Proceedings of SPIE*, pp. 68–79, San Diego, Calif, USA, July 1995.

[3] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," *Applied and Computational Harmonic Analysis*, vol. 5, no. 3, pp. 332–369, 1998.

[4] A. Gouze, M. Antonini, and M. Barlaud, "Quincunx lifting scheme for lossy image compression," in *Proceedings of IEEE International Conference on Image Processing (ICIP '00)*, vol. 1, pp. 665–668, Vancouver, BC, Canada, September 2000.

[5] C. Guillemot, A. E. Cetin, and R. Ansari, "*M*-channel nonrectangular wavelet representation for 2-D signals: basis for quincunx sampled signals," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '91)*, vol. 4, pp. 2813–2816, Toronto, Ontario, Canada, April 1991.

[6] R. Ansari and C.-L. Lau, "Two-dimensional IIR filters for exact reconstruction in tree-structured sub-band decomposition," *Electronics Letters*, vol. 23, no. 12, pp. 633–634, 1987.

[7] R. Ansari, A. E. Cetin, and S. H. Lee, "Subband coding of images using nonrectangular filter banks," in *The 32nd Annual International Technical Symposium: Applications of Digital Signal Processing*, vol. 974 of *Proceedings of SPIE*, p. 315, San Diego, Calif, USA, August 1988.

[8] A. Gouze, M. Antonini, M. Barlaud, and B. Macq, "Design of signal-adapted multidimensional lifting scheme for lossy coding," *IEEE Transactions on Image Processing*, vol. 13, no. 12, pp. 1589–1603, 2004.

[9] W. Trappe and K. J. R. Liu, "Adaptivity in the lifting scheme," in *Proceedings of the 33rd Annual Conference on Information Sciences and Systems*, pp. 950–955, Baltimore, Md, USA, March 1999.

[10] A. Benazza-Benyahia and J.-C. Pesquet, "Progressive and lossless image coding using optimized nonlinear subband decompositions," in *Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '99)*, vol. 2, pp. 761–765, Antalya, Turkey, June 1999.

[11] Ö. N. Gerek and A. E. Çetin, "Adaptive polyphase subband decomposition structures for image compression," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1649–1660, 2000.

[12] R. L. Claypoole, G. M. Davis, W. Sweldens, and R. G. Baraniuk, "Nonlinear wavelet transforms, for image coding via lifting," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1449–1459, 2003.

[13] G. Piella and H. J. A. M. Heijmans, "Adaptive lifting schemes with perfect reconstruction," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1620–1630, 2002.

[14] G. Piella, B. Pesquet-Popescu, and H. Heijmans, "Adaptive update lifting with a decision rule based on derivative filters," *IEEE Signal Processing Letters*, vol. 9, no. 10, pp. 329–332, 2002.

[15] J. Solé and P. Salembier, "Adaptive discrete generalized lifting for lossless compression," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 3, pp. 57–60, Montreal, Quebec, Canada, May 2004.

[16] D. S. Taubman, "Adaptive, non-separable lifting transforms for image compression," in *Proceedings of IEEE International Conference on Image Processing (ICIP '99)*, vol. 3, pp. 772–776, Kobe, Japan, October 1999.

[17] N. V. Boulgouris, D. Tzovaras, and M. G. Strintzis, "Lossless image compression based on optimal prediction, adaptive lifting, and conditional arithmetic coding," *IEEE Transactions on Image Processing*, vol. 10, no. 1, pp. 1–14, 2001.

[18] Ö. N. Gerek and A. E. Çetin, "A 2-D orientation-adaptive prediction filter in lifting structures for image coding," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 106–111, 2006.

[19] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic, Boston, Mass, USA, 2002.

[20] A. Benazza-Benyahia, J.-C. Pesquet, and M. Hamdi, "Vector-lifting schemes for lossless coding and progressive archival of multispectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2011–2024, 2002.

[21] A. Benazza-Benyahia, J.-C. Pesquet, and H. Masmoudi, "Vector-lifting scheme for lossless compression of quincunx sampled multispectral images," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02)*, p. 3, Toronto, Ontario, Canada, June 2002.

[22] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[23] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions of Image Processing*, vol. 1, no. 2, pp. 205–220, 1992.

[24] K. Sharifi and A. Leron-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 1, pp. 52–56, 1995.

[25] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, no. 5, pp. 676–683, 1968.

[26] J. Hattay, A. Benazza-Benyahia, and J.-C. Pesquet, "Adaptive lifting schemes using variable-size block segmentation," in *Proceedings of International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS '04)*, pp. 311–318, Brussels, Belgium, August-September 2004.

[27] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.

[28] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1303–1310, 1996.

[29] D. S. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, 2000.

[30] J. Hattay, A. Benazza-Benyahia, and J.-C. Pesquet, "Multi-component image compression by an efficient coder based on vector lifting structures," in *Proceedings of the 12th IEEE International Conference on Electronics, Circuits and Systems (ICECS '05)*, Gammarth, Tunisia, December 2005.

[31] S. R. Tate, "Band ordering in lossless compression of multispectral images," *IEEE Transactions on Computers*, vol. 46, no. 4, pp. 477–483, 1997.

[32] P. Hao and Q. Shi, "Reversible integer KLT for progressive-to-lossless compression of multiple component images," in *Proceedings of IEEE International Conference on Image Processing (ICIP '03)*, vol. 1, pp. 633–636, Barcelona, Spain, September 2003.

[33] H. Masmoudi, A. Benazza-Benyahia, and J.-C. Pesquet, "Block-based adaptive lifting schemes for multiband image compression," in *Wavelet Applications in Industrial Processing*, vol. 5266 of *Proceedings of SPIE*, pp. 118–128, Providence, RI, USA, October 2003.

[34] J. Hattay, A. Benazza-Benyahia, and J.-C. Pesquet, "Adaptive lifting for multicomponent image coding through quadtree partitioning," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 2, pp. 213–216, Philadelphia, Pa, USA, March 2005.