TWO-DIMENSIONAL NON SEPARABLE ADAPTIVE LIFTING SCHEME FOR STILL AND STEREO IMAGE CODING

M. Kaaniche[1,2], J.- C. Pesquet[3], A. Benazza-Benyahia[2] and B. Pesquet-Popescu [1]

[1] Telecom ParisTech,
Signal and Image Proc. Dept.
37-39, rue Dareau, 75014 Paris, France
kaaniche@telecom-paristech.fr
pesquet@telecom-paristech.fr

[2] SUP'COM, URISA
Cité Technologique des
Communications, 2083, Tunisia
benazza.amel@supcom.rnu.tn

[3] Université Paris-Est,
IGM and UMR-CNRS 8049,
Champs-sur-Marne, 77454
Marne-la-Vallée, France
pesquet@univ-mlv.fr

ABSTRACT

Many existing works related to lossy-to-lossless image compression are based on the lifting concept. However, it has been observed that the separable lifting scheme structure presents some limitations because of the separable processing performed along the image lines and columns. In this paper, we propose to use a 2D non separable lifting scheme decomposition that enables progressive reconstruction and exact decoding of images. More precisely, we focus on the optimization of *all* the involved decomposition operators. In this respect, we design the prediction filters by minimizing the variance of the detail signals. Concerning the update filters, we propose a new optimization criterion which aims at reducing the inherent aliasing artefacts. Simulations carried out on still and stereo images show the benefits which can be drawn from the proposed optimization of the lifting operators.

Index Terms— lossy-to-lossless image compression, lifting schemes, separable transforms, non separable transforms, adaptive transforms, multiresolution analysis, stereo coding.

1. INTRODUCTION

The discrete wavelet transform has been studied and applied extensively in many image processing fields such as denoising [1] and compression [2]. In this respect, the second generation of wavelets provides very efficient transforms, based on the concept of Lifting Scheme (LS) developed by Sweldens [3]. It was shown that interesting properties are offered by such structures. Indeed, lifting schemes are suitable tools for scalable reconstruction, which is a key issue for telebrowsing applications. Besides, they guarantee a lossy-to-lossless reconstruction required in some specific applications such as medical imaging for which any distortion in the decoded image may lead to an erroneous interpretation of the image [4, 5]. A generic LS applied to a 1D signal consists of three modules referred as *split*, *predict* and *update*. In the first step, the even samples are separated from the odd ones. Then, each sample of one of the two resulting subsets (say the even one) is predicted from the odd samples and a prediction error or detail coefficient is computed. Finally, the update step generates a coarser approximation of the initial signal thanks to a smoothing of the odd samples using the detail coefficients. Generally, for 2D signals, the LS is handled in a separable way by cascading 1D LS along the horizontal direction, then along the vertical one. However, this separate processing limits the number of freedom degrees of the involved operators [6]. Furthermore, separable LS is not efficient to cope with characteristics of edges which are neither horizontal nor vertical. In addition to still monocular images, LS have been extensively used for stereo image coding [7, 8]. Indeed, it can be noticed that most of the existing works rely on disparity compensation techniques. The first step in this approach consists of estimating the disparity map. Then, one image is considered as a reference image and the other is predicted in order to generate a prediction error referred to as a residual image. Finally, the disparity field, the reference image and the residual one are encoded. In this context, Moellenhoff and Maier [9] analyzed the characteristics of the residual image and proved that such images have properties different from natural images. This suggests that transforms working well for natural images may not be as well-suited for residual images. These shortcomings have motivated the development of 2D Non Separable LS (NSLS) in order to offer more flexibility in the design of the prediction filter. Indeed, instead of using samples from the same rows (resp. columns) while processing the image along the lines (resp. columns), 2D NSLS provide more choices in the selection of the samples by using horizontal, vertical and oblique directions [6]. Furthermore, operators in an NSLS can satisfy some appealing properties (e.g. orthogonality) unlike those related to a separable LS [10]. The orthogonality property may be interesting in image coding since the mean square error in the wavelet domain and in the spatial domain are the same. Also, in [11], it was shown that 2D NSLS structures outperform separable ones due to the reduction of the rounding effects. Moreover, the performance of these LS can be improved by exploiting the characteristics of the input image. However, to the best of our knowledge, most existing works have focused on the optimization of the prediction operators. In this context, Gerek and *al.* [6] proposed to consider three direction angles of prediction ($0°$, $45°$ and $135°$) and they selected the orientation which leads to the minimum prediction error. In [12], the authors considered a non separable quincunx LS and proposed to optimize the predictors based on an entropy criterion. However, it can be noticed that there are few works which have discussed the problem of the update filter. Among these works, in [13], the update operator of a separable LS is adaptively computed thanks to a non linear decision rule using the local gradient information. Another alternative is to adapt the update so that the reconstruction error is minimized when the detail coefficients are canceled [14]. In this paper, we aim at fully exploiting the flexibility of a NSLS through an optimization of *all* the involved operators. One of the main contributions of this work relies on the optimization of the update filter by using a criterion which allows us to reduce the aliasing effects. The outline of the paper is as follows. In Section 2, the 2D NSLS and its link with some existing 1D LS structures are presented. The novel optimiza-

tion method of both the prediction and update filters is discussed in Section 3. Finally, in Section 4, experimental results are given and some conclusions are drawn in Section 5.

2. 2D NON SEPARABLE LS STRUCTURE

2.1. Principle

Let x denote the digital image to be coded. At each resolution level j and each pixel location (m, n), the approximation coefficient $x_j(m, n)$ has 4 polyphase components $x_{0,j}(m, n) = x_j(2m, 2n)$, $x_{1,j}(m, n) = x_j(2m, 2n + 1)$, $x_{2,j}(m, n) = x_j(2m + 1, 2n)$, and $x_{3,j}(m, n) = x_j(2m + 1, 2n + 1)$. Without loss of generality, we assume that the polyphase components are the input coefficients of the 2D NSLS depicted in Fig. 1, where $\mathbf{P}_j^{(HH)}$, $\mathbf{P}_j^{(LH)}$, $\mathbf{P}_j^{(HL)}$ and $\mathbf{U}_j$ represent the four analysis filters employed to generate the detail coefficients: $x_{j+1}^{(HH)}$ oriented diagonally, $x_{j+1}^{(LH)}$ oriented vertically, $x_{j+1}^{(HL)}$ oriented horizontally, and the approximation coefficients x_{j+1}. It is easy to derive the expressions of the resulting coefficients
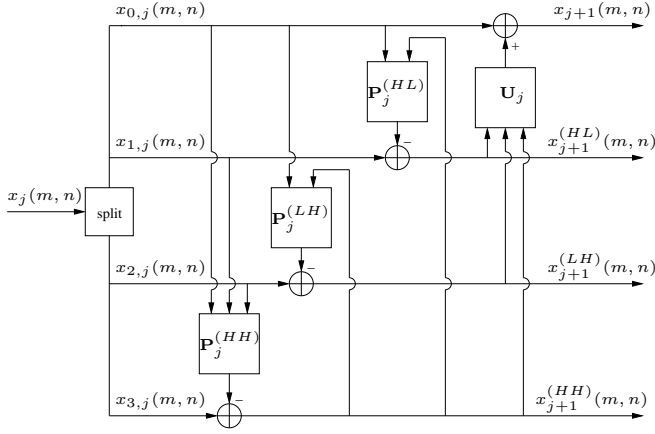

Fig. 1. NSLS decomposition structure.

in the 2D z-transform domain.[1] Indeed, it is straightforward to show that the z-transforms of the output coefficients are as follows:

$$X_{j+1}^{(HH)}(z_1, z_2) = X_{3,j}(z_1, z_2) - \lfloor P_{0,j}^{(HH)}(z_1, z_2)X_{0,j}(z_1, z_2) + P_{1,j}^{(HH)}(z_1, z_2)X_{1,j}(z_1, z_2) + P_{2,j}^{(HH)}(z_1, z_2)X_{2,j}(z_1, z_2) \rfloor, \quad (1)$$

$$X_{j+1}^{(LH)}(z_1, z_2) = X_{2,j}(z_1, z_2) - \lfloor P_{0,j}^{(LH)}(z_1, z_2)X_{0,j}(z_1, z_2) + P_{1,j}^{(LH)}(z_1, z_2)X_{j+1}^{(HH)}(z_1, z_2) \rfloor, \quad (2)$$

$$X_{j+1}^{(HL)}(z_1, z_2) = X_{1,j}(z_1, z_2) - \lfloor P_{0,j}^{(HL)}(z_1, z_2)X_{0,j}(z_1, z_2) + P_{1,j}^{(HL)}(z_1, z_2)X_{j+1}^{(HH)}(z_1, z_2) \rfloor, \quad (3)$$

$$X_{j+1}(z_1, z_2) = X_{0,j}(z_1, z_2) + \lfloor U_j^{(0)}(z_1, z_2)X_{j+1}^{(HL)}(z_1, z_2) + U_j^{(1)}(z_1, z_2)X_{j+1}^{(LH)}(z_1, z_2) + U_j^{(2)}(z_1, z_2)X_{j+1}^{(HH)}(z_1, z_2) \rfloor \quad (4)$$

---

[1]The z-transform of a signal x will be denoted in capital letters by X.

where $\forall i \in \{0, 1, 2\}$ and $o \in \{HH, HL, LH\}$,

$$P_{i,j}^{(o)}(z_1, z_2) = \sum_{(k,l) \in \mathcal{P}_{i,j}^{(o)}} p_{i,j}^{(o)}(k, l)z_1^{-k}z_2^{-l}, \quad (5)$$

$$U_j^{(i)}(z_1, z_2) = \sum_{(k,l) \in \mathcal{U}_j^{(i)}} u_j^{(i)}(k, l)z_1^{-k}z_2^{-l}. \quad (6)$$

The set $\mathcal{P}_{i,j}^{(o)}$ (resp. $\mathcal{U}_j^{(i)}$) and the coefficients $p_{i,j}^{(o)}(k, l)$ (resp. $u_j^{(i)}(k, l)$) denote the support and the weights of the three prediction filters (resp. of the update filter). Once the considered NSLS structure has been defined, it may appear instructive to see how it is related to some commonly used 1D LS structures.

2.2. Links with conventional 1D LS

It can be checked that the conventional 5/3 transform and the Haar one are particular cases of the structure illustrated in Fig. 1. For example, consider the 5/3 transform which was selected for the lossless mode of the JPEG2000 standard [5]. After applying the 5/3 decomposition to the lines then to the columns, we deduce that

$$P_{0,j}^{(HH)}(z_1, z_2) = -\frac{1}{4}(1 + z_1 + z_2 + z_1z_2),$$

$$P_{1,j}^{(HH)}(z_1, z_2) = \frac{1}{2}(1 + z_1), P_{2,j}^{(HH)}(z_1, z_2) = \frac{1}{2}(1 + z_2), \quad (7)$$

$$P_{0,j}^{(LH)}(z_1, z_2) = \frac{1}{2}(1 + z_1), P_{1,j}^{(LH)}(z_1, z_2) = -\frac{1}{4}(1 + z_2^{-1}), \quad (8)$$

$$P_{0,j}^{(HL)}(z_1, z_2) = \frac{1}{2}(1 + z_2), P_{1,j}^{(HL)}(z_1, z_2) = -\frac{1}{4}(1 + z_1^{-1}), \quad (9)$$

$$U_j^{(0)}(z_1, z_2) = \frac{1}{4}(1 + z_1^{-1}), U_j^{(1)}(z_1, z_2) = \frac{1}{4}(1 + z_2^{-1}),$$

$$U_j^{(2)}(z_1, z_2) = -\frac{1}{16}(1 + z_1^{-1} + z_2^{-1} + z_1^{-1}z_2^{-1}). \quad (10)$$

This shows that some widely used separable LS can thus be put under the form of a 2D-NSLS. This will be shown to facilitate the derivation of adaptive extensions of the associated operators. Indeed, in a coding framework, the compactness of any LS-based multiresolution representation clearly depends on the choice of the prediction and update operators. In the next section, we address the issue of an optimal design of the NSLS operators.

3. PROPOSED OPTIMIZATION METHOD

3.1. Optimization of the predictors

As the detail coefficients are defined as prediction errors, the prediction operators can be optimized so as to minimize the variance of the coefficients at each resolution level. If the rounding operators are omitted, it is readily shown that the minimum variance predictors must satisfy the well-known Yule-Walker equations. For example, for the prediction vector $\mathbf{P}_j^{(HH)}$, the normal equations read

$$\mathsf{E}[\tilde{\mathbf{x}}_j(m, n)\tilde{\mathbf{x}}_j(m, n)^\top]\mathbf{P}_j^{(HH)} = \mathsf{E}[x_{3,j}(m, n)\tilde{\mathbf{x}}_j(m, n)] \quad (11)$$

where
• $\mathbf{P}_j^{(HH)} = (\mathbf{P}_{0,j}^{(HH)}, \mathbf{P}_{1,j}^{(HH)}, \mathbf{P}_{2,j}^{(HH)})$ is the prediction vector where, for each $i \in \{0, 1, 2\}$, $\mathbf{P}_{i,j}^{(HH)} = \left(p_{i,j}^{(HH)}(k, l)\right)_{(k,l) \in \mathcal{P}_{i,j}^{(HH)}}$,

- $\tilde{\mathbf{x}}_j(m,n) = (\mathbf{x}_{0,j}(m,n), \mathbf{x}_{1,j}(m,n), \mathbf{x}_{2,j}(m,n))$ is the reference vector with $\mathbf{x}_{i,j}(m,n) = \left(x_{i,j}(m-k, n-l) \right)_{(k,l) \in \mathcal{P}_{i,j}^{(HH)}}$.

The other optimal prediction filters $\mathbf{P}_j^{(HL)}$ and $\mathbf{P}_j^{(LH)}$ are obtained in a similar way.

3.2. Optimization of the update operator

Firstly, it can be noticed that the transfer function of the low-pass filter relating x_j to the undecimated version of x_{j+1} can be expressed as

$$F_j(z_1, z_2) = 1 + \sum_{i \in \{0,1,2\}} \sum_{(k,l) \in \mathcal{U}_j^{(i)}} u_j^{(i)}(k,l) z_1^{-2k} z_2^{-2l} A_j^{(i)}(z_1, z_2)$$

where $A_j^{(i)}(z_1, z_2)$ is a transfer function which depends on the prediction coefficients. Since F_j aims at computing a smooth version of the original image, we propose to design the update filter in order to reduce the aliasing effects. More precisely, we adopt a new criterion $\mathcal{J}$ which accounts for the difference between the output of the F_j filter and the output of an ideal filter

$$\mathcal{J}(u_j) = \int_{-\pi}^{+\pi} \left| F_j(e^{\iota\omega_1}, e^{\iota\omega_2}) \right. \\ \left. - H(e^{\iota\omega_1}, e^{\iota\omega_2}) \right|^2 S_{x_j}(\omega_1, \omega_2) d\omega_1 d\omega_2 \quad (12)$$

where S_{x_j} is the power spectrum density of x_j and H denotes the transfer function of the ideal low-pass filter (h will subsequently designate its impulse response). It must be emphasized that this criterion exploits the characteristics of the input signal, so making the method adaptive. In this respect, it departs from criteria classically employed for filter bank design. By cancelling the gradient of $\mathcal{J}$, the optimal update weights should satisfy, for each $(k_0, l_0) \in \mathcal{U}_j^{(i')}$ and $i' \in \{0, 1, 2\}$, the following condition:

$$\sum_{i \in \{0,1,2\}} \sum_{(k,l) \in \mathcal{U}_j^{(i)}} u_j^{(i)}(k,l) \sum_{(m,n)} \sum_{(m',n')} a_j^{(i)}(m,n) a_j^{(i')}(m',n')$$
$$\times \gamma_{x_j}(m'-m+2k_0-2k, n'-n+2l_0-2l) =$$
$$\sum_{(m,n)} \sum_{(m',n')} h(m,n) a_j^{(i')}(m',n') \gamma_{x_j}(m'-m+2k_0, n'-n+2l_0)$$
$$- \sum_{(m',n')} a_j^{(i')}(m',n') \gamma_{x_j}(m'+2k_0, n'+2l_0) \quad (13)$$

where γ_{x_j} is the autocorrelation function of the field x_j. Consequently, the optimal update weights are solutions of a *linear* system of equations. Fig. 2 shows the magnitude of the frequency responses of the F_0 filter for the 5/3 transform, obtained before and after our optimization method on the residual image generated from the stereo pair "apple". It can be clearly seen that the optimal filter is quite different from that obtained when no optimization is performed.

4. EXPERIMENTAL RESULTS

Experiments were firstly carried out on test images by considering a separable decomposition and its corresponding 2D-NSLS. Then, the gain related to the optimization of the NSLS operators was evaluated.

More precisely, we considered the 5/3 separable transform (denoted by "DEC-5/3-SEP") which was found to be a very effective for
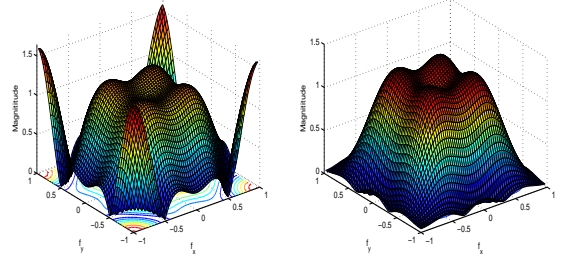

Fig. 2. Frequency responses of the filter F_0: (left figure) before optimization, (right figure) after optimization

lossy-to-lossless image coding [5]. The number of resolution levels was set to 2. In the following, the corresponding NSLS decomposition will be designated by "DEC-5/3-NON SEP". It is worth recalling that "DEC-5/3-NON SEP" is non adaptive since its underlying operators are given by Equations (7)-(10). The last set of experiments was devoted to the optimization of the operators by using the equivalent *non separable* decomposition, their supports being preserved.

- First, the prediction filters were optimized while the update filters were kept. In the following, this method will be designated by "DEC-5/3-NON SEP-OPT1".
- The second variant (denoted by "DEC-5/3-NON SEP-OPT2") consists of optimizing *both* the predictor and the update. Hence, we used the method proposed in the previous section for designing the update filter.

The performance of the considered methods has been assessed on still images[2] and also on stereo images. Fig. 3 gives the variations of the PSNR versus the bitrate for the image "straw". It is worth pointing out that "DEC-5/3-NON SEP-OPT2" outperforms the "DEC-5/3-NON SEP-OPT1" by 0.1-0.35 dB. Fig. 5 displays a zoom applied on the reconstructed images at 0.6 bpp. It can be observed that the optimization of the prediction and update filters improves the quality of reconstruction. In addition, we have tested these methods in a lossless coding context and the related final bitrates are given in Tab. 1. Slight improvements are obtained when our optimization method is employed for a lossless coding scheme. The second part of the experiments is concerned with stereo images. For this reason, we also proposed to evaluate our method for residual images. Fig. 4 illustrates the evolution of the PSNR versus the bitrate of the residual image generated from the "apple" stereo image downloaded from [3]. An improvement of 0.1-0.4 dB is obtained by optimizing the update filter. The gain becomes more important (up to 0.65 dB) when compared with the non-optimized 5/3 transform. This confirms the effectiveness of our method in terms of quality of reconstruction.

Table 1. Performance of the lossless decompositions in terms of bitrate (bpp) using JPEG2000.

Image	DEC-5/3-SEP	DEC-5/3-NONSEP	DEC-5/3-NONSEP-OPT1	DEC-5/3-NONSEP-OPT2
spot5-1	3.775	3.777	3.727	3.725
spot5-2	3.780	3.781	3.728	3.727
straw	6.346	6.349	6.343	6.341

[2]taken from the URL http://sipi.usc.edu/database
[3]http://vasc.ri.cmu.edu/idb/html/stereo/index.html
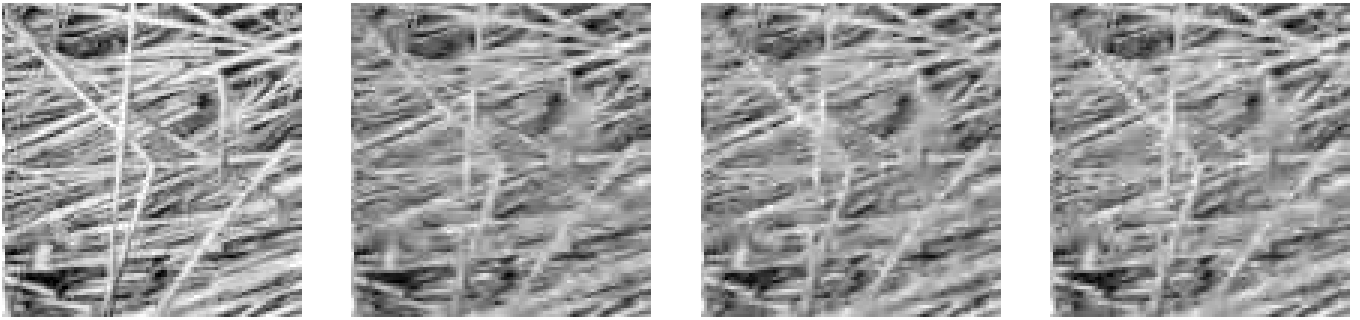
(a) Original image   (b) PSNR=23.67 dB, SSIM=0.808   (c) PSNR=24.03 dB, SSIM=0.824   (d) PSNR=24.34 dB, SSIM=0.826

Fig. 5. Zoom applied on the reconstructed "straw" image using (b) "DEC-5/3-SEP" (c) "DEC-5/3-NON SEP-OPT1" (d) "DEC-5/3-NON SEP-OPT2".
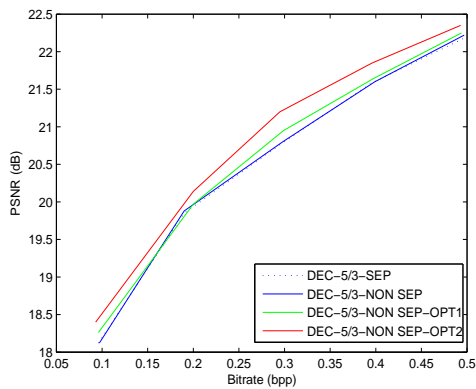

Fig. 3. PSNR (in dB) versus the bitrate (bpp) after JPEG2000 progressive encoding for the "straw" image.
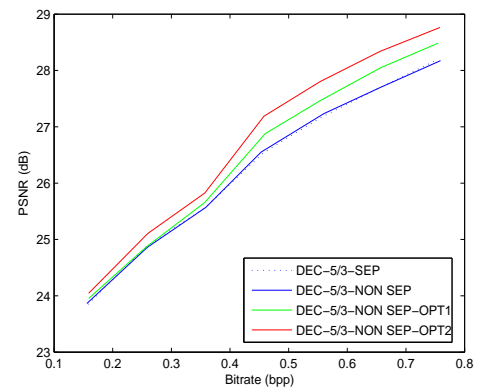

Fig. 4. PSNR (in dB) versus the bitrate (bpp) after JPEG2000 progressive encoding of the residual image of the SI pair "apple".

5. CONCLUSIONS

In this paper, we have exploited the flexibility offered by non separable lifting schemes to perform their optimization. A new criterion has been presented for the optimization of the update filter in this context. The proposed method adapts the filter to the contents of the input image while ensuring perfect reconstruction. Experimental results, carried out on still images and residual images of stereo pairs, have illustrated the benefits of optimizing both the prediction and the update filters. In future work, we plan to extend this optimization method to the vector lifting scheme recently presented for stereo image coding [7].

6. REFERENCES

[1] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425-455, September 1994.

[2] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 205-220, April 1992.

[3] W. Sweldens, "The lifting scheme: A custom design of biorthogonal wavelets," *J. Appl. Comput. Harmonic Anal.*, vol. 3, no. 2, pp. 186-200, 1996.

[4] A. R. Calderbank, I. Daubechies, W. Sweldens and B.-L. Yeo, "Wavelet transforms that map integers to integers," *J. Appl. Comput. Harmonic Anal.*, vol. 5, pp. 332-369, 1998.

[5] D. Taubman and M. Marcellin, *JPEG2000: Image compression fundamentals, standards and practice*, Springer Verlag, 2001.

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

[7] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu and J.-C. Pesquet, "Vector lifting schemes for stereo image coding," *to appear in IEEE Transactions on Image Processing*, November 2009.

[8] N. V. Boulgouris and M. G. Strintzis, "A family of wavelet-based stereo image coders," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 898-903, October 2002.

[9] M. S. Moellenhoff and M. W. Maier, "Characteristics of disparity-compensated stereo image pair residuals," *Signal Processing, Image Communications*, vol. 14, no. 1-2, pp. 55-69, November 1998.

[10] S. Sezen and A. Ertuzun, "2D Four-Channel Perfect Reconstruction Filter Bank Realized with the 2D Lattice Filter Structure," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1-16, January 2006.

[11] S. Chokchaitam, "A non-separable two-dimensional LWT for an image compression and its theoretical analysis", *Thammasat International journal of science and technology*, vol. 9, no. 1, pp. 35-43, January-March 2004.

[12] A. Benazza-Benyahia, J.-C. Pesquet, J. Hattay and H. Masmoudi, "Block-based adaptive vector lifting schemes for multichannel image coding coding," *EURASIP Journal on Image and Video Processing*, vol. 2007, 10 pages, January 2007.

[13] 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, October 2002.

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