NONLINEAR BANK FILTERS ADAPTED FOR PROGRESSIVE AND EXACT IMAGE RECONSTRUCTION

A. Benazza-Benyahia [(1)], *J.-C. Pesquet* [(2)], *A. Amara* [(1)]
(1) Ecole Supérieure des Communications de Tunis,
3.5 Km Route de Raoued, 2083 Cité El-Ghazala, TUNISIA,
Tel: +216 1 856440; fax: +216 1 856829
e-mail: ben.yahia@planet.tn
(2) LSS (CNRS/UPS)
Supélec, Plateau de Moulon, 91192, Gif sur Yvette Cédex, FRANCE
e-mail: Jean-Christophe.Pesquet@univ-mlv.fr

ABSTRACT

In this paper, an extension of the recently proposed nonlinear subband decomposition schemes with perfect reconstruction is investigated. Such generalized scheme has the advantage of exploiting efficiently the redundancies contained in the images. Simulation tests performed on a great number of natural images show that the proposed method is suitable for lossless and progressive transmission.

1 INTRODUCTION

The use of progressive transmission techniques is recommended for images databases telebrowsing applications. From a coarse version of the image, the user can quickly recover the original image by a gradual refinement of the approximate versions. The amount of information corresponding to these coarse versions is lower than the information required for the direct transmission of the image. This results in a gain in terms of bit rate, especially when the user does proceed to a complete download of the image.

The wavelet decompositions [1] provide very compelling representations for the progressive coding of the images thanks to their hierarchical properties. However, the coders corresponding to linear decompositions are generally lossy since the coefficients must be rounded/quantized in the binary transcription step. However, for some application such as telemedicine or satellite imaging, the least distortion in the reconstructed signal constitutes a risk (even of minor importance) of making erroneous diagnosis. In the context of such applications, the greatest care must be taken to ensure an exact coding. Therefore, the multiresolution decomposition must guarantee a perfect reconstruction. To this respect, new methods of nonlinear subband decompositions, based on "lifting" schemes, have been recently introduced [2, 3, 4].

The objective of our study is to generalize such nonlinear decompositions in order to exploit efficiently the redundancies contained in the signal. When the image includes complex features, it is proposed to adaptively select an appropriate "best decomposition" so as to achieve better compression performances.

2 PROPOSED NONLINEAR DECOMPOSITION SCHEME

The 1D nonlinear subband decomposition structure is depicted in Fig. 1 [4]. We initialize the decomposition process by setting $a_0(n) = x(n)$, where $x(n)$ denotes the signal to be coded. By an appropriate choice of the operators $\mathcal{H}$ and $\mathcal{G}$, the coefficients $a_j(n)$ may be viewed as the approximation coefficients of the signal at resolution level j whereas $d_j(n)$ are associated with the details lost when passing from resolution level j to the next coarser one $(j+1)$. The two subband signal $a_j(n)$ and $d_j(n)$ are given by:

$$\left\{ \begin{array}{lcl} d_{j+1}(n) & = & \mathcal{A}_2[(a_j(2n+1))_{n\in\mathbb{Z}}] - \mathcal{G}[\mathcal{A}_1[\mathbf{a}_j(2n)]] \\ a_{j+1}(n) & = & \mathcal{A}_1[(a_j(2n))_{n\in\mathbb{Z}}] + \mathcal{H}[\mathbf{d}_{j+1}(n)] \end{array} \right. , \quad (1)$$

where $\mathbf{a}_j(n) \triangleq \big(a_j(2(n+k))\big)_{-N_1\leq k\leq N_2}$ and $\mathbf{d}_j(n) \triangleq (d_j(n+k))_{-N_1'\leq k\leq N_2'}$. Such nonlinear extensions of linear filter banks with critical subsampling allow an exact reconstruction provided that $\mathcal{A}_1$ and $\mathcal{A}_2$ are one-to-one mappings [4]. The approximation signal is recursively processed over j_m resolution levels and, in this way, a pyramid representation is built. Extension to 2D signal can be easily carried out by defining polyphase components of the input image. For example, rectangular or quincux subsampling structures can be used. For the sake of simplicity, only separable decompositions (rectangular subsampling structures) are considered in the sequel.

An important merit of this decomposition structure is to provide a unifiying framework for the main nonlinear

decomposition schemes proposed in the last years [5]. At this point, it is interesting to mention the main promising schemes frequently used in this field. The first class is related to the nonexpansive Laplacian-like pyramids [6] and more precisely the Reduced Laplacian Pyramid (RLP) and the Minimum Entropy Pyramid (MEP). The second class belongs to second generation wavelets that map integers to integers, based on the concept of lifting, developed by Sweldens [7] as well as other researchers. It is worth pointing out that in the work of Sweldens [7], the operators $\mathcal{H}$ and $\mathcal{G}$ are restricted to be FIR linear filters followed by a rounding step and $\mathcal{A}_1 = \mathcal{A}_2 = \mathrm{Id}$. Several examples of such operators have been tabulated in [3]. In the following, we will denote by $c_{N,\tilde{N}}$ the wavelet transforms that map integers to integers where the numbers $N, \tilde{N}$ correspond repectively to the number of vanishing moments of the analyzing (resp. synthesizing) high-pass filters. By varying $(N, \tilde{N})$, it is possible to obtain a family of reversible wavelet tranforms that map integers to integers.
The third class of decomposition (denoted $\mathrm{T}(\alpha)$) can be written in the following manner:

$$\begin{aligned} d_{j+1}(n) &= a_j(2n+1) - \lfloor \mathbf{b}^T \mathbf{a}_j(n) \rceil \\ a_{j+1}(n) &= a_j(2n) + \lfloor 0.25(d_{j+1}(n) + d_{j+1}(n-1)) \rceil \end{aligned} \quad (2)$$

where $\lfloor \cdot \rceil$ designates the rounding operation, $\mathbf{a}_j(n) = [a_j(2n-1), a_j(2n), a_j(2n+2), a_j(2n+4)]^T$ and $\mathbf{b} = [\frac{1-\alpha}{2}, \frac{\alpha}{2}, \frac{1+\alpha}{4}, \frac{1-\alpha}{4}]^T$. The parameter α of the prediction vector $\mathbf{b}$ can be chosen in order to decrease the entropy of the one-stage decomposition $\mathrm{T}(\alpha)$. The main difference with the transformation $c_{N,\tilde{N}}$ is the anti-causal terms $a_j(2n+2)$, $a_j(2n+4)$ for the prediction of the odd sample $a_j(2n+1)$. Decomposition $\mathrm{T}(\alpha)$ is similar to the well-known S+P transform [2] because of the anti-causal terms. Note that $\mathrm{T}(\alpha)$ does not exactly fit in the structure of Fig. 1 (due to the use of $a_j(2n-1)$ for the prediction) but it can be easily included in an extended class of nonlinear filter banks with perfect reconstruction [4].

When the signal $x(n)$ contains detailed and complex features corresponding to high frequency components, it is preferable to adapt the considered multiresolution decomposition to the signal. So, inspired by the compression techniques based on wavelet packets decompositions [8], [9], we propose an extension of flexible subband decompositions to the nonlinear case. This extension consists in analyzing *both* the approximation and detail signals at each level of a pyramidal decomposition. At each resolution level j, the sequences of generated coefficients $(a_{j,m}(n))_n$ with $m \in \{0, \ldots, 2^j - 1\}$ can be associated with the nodes of a binary tree. The "father-node" $(a_{j,m}(n))_n$ is iteratively decomposed into the two "children-nodes" $(a_{j+1,2m}(n))_n$, $(a_{j+1,2m+1}(n))_n$ using nonlinear filter banks. We designate such a new Extended NonLinear subband decomposition by the abbreviation ENL. Obviously, in the case of 2D signals, a full quadtree is generated.

3 "BEST" DECOMPOSITION SELECTION

Similarly to wavelet packets decompositions, it is not necessary to decompose systematically each node (j,m). Indeed, for every partition $\mathcal{P}$ of $[0,1[$ in intervals $I_{j,m} = [m2^{-j}, (m+1)2^{-j}[$, the sequences $\{(a_{j,m}(n))_n,\ (j,m) \mid I_{j,m} \in \mathcal{P}\}$ provide a complete representation of the input signal. So, it seems appropriate to exploit such flexibility of the analysis in order to adapt the decomposition to the redundancies contained in the original image. More precisely, we adopt the strategy of Wickerhauser [8] to select the "best" representation by pruning the full tree. The "best" selection consists in choosing the optimal subtree with respect to a predefined criterion. The usual criterion used by Wickerhauser [8] in the context of lossy compression is the "entropy" $H_W[\cdot]$ of the energy distribution of the decomposition coefficients:

$$H_W[a_{j,m}] = -\sum_i \frac{\mid a_{j,m}(i) \mid^2}{A_{j,m}} \log_2 \left(\frac{\mid a_{j,m}(i) \mid^2}{A_{j,m}} \right), \quad (3)$$

where $A_{j,m} \stackrel{\triangle}{=} \sum_i \mid a_{j,m}(i) \mid^2$ (with the convention $0 \log_2 0 = 0$). It is important to point out that strictly speaking, this is not an information theory entropy. Applying this entropy, the aim is to select the representation which leads to the best energy concentration. However, in lossless coding, the objective is somehow different: we look for a compact representation with exact reconstruction. So, it is more appropriate to measure the compactness of the hierarchical representation in terms of Shannon "entropy":

$$H_S[a_{j,m}] = -\sum_i P[a_{j,m}(i)] \log_2(P[a_{j,m}(i)]). \quad (4)$$

where $P[u]$ denotes the probability of occurence of intensity level u. The choice of such a criterion is computationnally simple and allows a good estimation of the bit-rate needed for coding.

In practice, once the complete tree of depth j_m, associated with the ENL is generated, the algorithm starts with the lowest level j_m by comparing the entropies of each parent with those of its children. More precisely, two "children"-nodes are deleted if the average of their entropies exceeds the entropy of their "father"-node, i.e. $H_S[a_{j-1,m}] < 0.5(H_S[a_{j,2m}] + H_S[a_{j,2m+1}])$.

In the opposite case, the signal is best represented by these "children"-nodes. Following such fusion rule, an unbalanced subtree of the inital complete tree is obtained. It is important to mention that a copy of the resulting quadtree must be available at the decoder. So, a side information must be sent, which consists of a an array of indexes describing at which level the next subband in the best basis is located. For a j_m-level decomposition, the levels are encoded with a very short integer

of at most $\log_2(j_m)$ bits. The resulting bitstream has to be scanned in the ascending way during the progressive decoding step.

We will denote by the acronym ENL* the decomposition ENL followed by an optimization of the tree-structured representation. It is worth noting that other criteria could be used for this optimization procedure. For example, in order to exploit the efficiency of embedded coders such as SPIHT [2], other pruning rules could be used as proposed recently in [11]. The resulting pyramid presents coherent spatial tree relationships which are well suited to tree-structured coding.

4 EXPERIMENTAL RESULTS

We have applied the considered decompositions to various monochrome images coded at 8 bpp, taken from the image database "SIMPA" provided by the French GdR-PRC ISIS (http://www-isis.enst.fr). Here, the effectiveness of the considered decompositions is measured by the zero-th order entropy H_S.
Table 1 indicates that the proposed ENL* scheme always leads to a significant decrease in entropy with respect to the "classical" wavelet-like decomposition (denoted by NL), reported in the literature, whatever the number of stages j_m is. Tables 2 and 3 provide the entropies of the pyramids resulting from the best known NL decompositions, considered as the state-of-the-art for nonlinear subband decompositions. For a fair comparison, we give also the performances of some monoresolution lossless coders. Indeed, the second line OLP corresponds to results obtained by applying an optimal third-order linear predictor based on the neighboring pixels $x(m,n-1), x(m-1,n-1)$ and $x(m-1,n)$. Results in the third line BJPEG are for the best among the eight JPEG predictors associated with the lossless mode of the JPEG standard. These tables show that there is no single transform that outperforms the other one for all the test images.

In fact, the best transform for each image strongly depends on the characteristics of the image. Furthermore, except for some images, multiresolution approaches give better results than monoresolution coders. Besides, thelatter ones do not allow progressive transmission. However, generalized and optimized decompositions ENL* provide potentially compact representations of the images thanks to the selection of the "best" representation. Fig. 2 shows an example of a pyramidal decomposition quite different from the one obtained with traditional wavelet-like transforms. The best representation corresponds to a "division of the frequency plane" in intervals of variable sizes, adapted to the content of the considered images.

References

[1] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, "Image coding using wavelet transform," IEEE Trans. on IP, vol. 1, pp. 205-220, April 1992.

[2] A. Said, W.A. Pearlman, "An image multiresolution representation for lossless and lossy compression," IEEE Trans. on IP , vol. 5, pp. 1303-1310, September 1996.

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

[4] F. Hampson, J.C. Pesquet, "M-band nonlinear subband decompositions with perfect reconstruction," IEEE Trans. on IP, vol. 7, pp. 1547-1560, November 1998.

[5] A. Benazza-Benyahia, J-. C. Pesquet, "A Unifying framework for lossless and progressive image coding," Proceedings of the International Symposium on Image/Video Communications over Fixed and Mobile Networks, Rabat, Morocco, April 2000, pp. 333-341.

[6] B. Aiazzi, L. Alparone, S. Baronti, "A Reduced Laplacian Pyramid for Lossless and Progressive Image Communication," IEEE on COM, vol. 44, pp. 18-22, January 1996.

[7] W. Sweldens, "The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions," Proceedings of SPIE, San-Diego, CA, USA, September 1995, vol. 2569, pp. 68-79.

[8] M.V. Wickerhauser, "INRIA lectures on wavelet packet algorithms," Proceedings ondelettes et paquets d'ondes, Rocquencourt, France, 17-21 juin 1991.

[9] K. Ramchandran, M. Vetterli, "Best wavelet backet bases in a rate-distortion sense," IEEE Trans. on IP, vol. 2, pp. 160-175, April 1993.

[10] A. Benazza-Benyahia, J-.C. Pesquet, "Progressive and Lossless Image Coding Using Optimized Nonlinear Subband Decompositions," Proceedings of the IEEE Nonlinear Signal Processing Conference, Antalya, Turkey, June 1999, pp. 761-765.

[11] H. Khalil, A. Jacquin, C. Podilchuk, "Constrained wavelet packets for tree-structured coding algorithms," Proceedings of the Data Compression Conference, Snowbird, Utah, USA, March 1999.

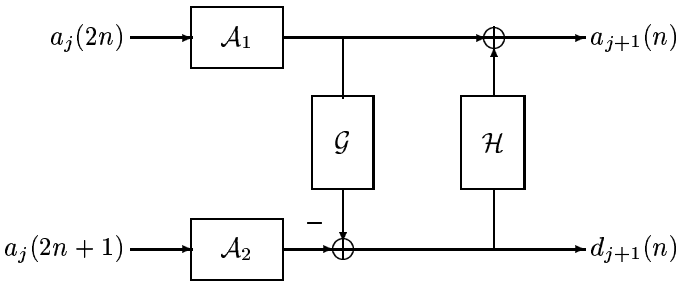

Figure 1: One stage of the nonlinear decomposition NL.

j_m	1	2	3
T(1.2) (NL)	6.6538	6.4063	6.3461
T(1.2) (ENL*)	6.6538	6.3975	6.3289
$c_{4,4}$ (NL)	6.6329	6.3773	6.3130
$c_{4,4}$ (ENL*)	**6.6329**	**6.3773**	**6.3077**
T(1.2) (NL)	4.9362	4.4266	4.3219
T(1.2) (ENL*)	4.9362	4.4266	4.3204
$c_{4,4}$ (NL)	4.9389	4.4403	4.3364
$c_{4,4}$ (ENL*)	**4.9365**	**4.4222**	**4.3160**

Table 1: Pyramid entropies H_S *versus* j_m for the images "Monkey" (first 4 lines), "Lena" (last 4 lines).

Image	angio	irm	echogr	muscle
Original	6.1702	5.4578	5.5732	7.4078
OLP	4.5731	4.5065	4.9398	**5.2800**
BJPEG	**3.9627**	**4.4761**	**4.1800**	5.3366
MEP (NL)	4.6472	4.4337	5.2486	5.4013
MEP (ENL*)	4.6956	4.4335	5.1479	5.3947
RLP (NL)	4.6956	4.4563	5.2028	5.3753
RLP (ENL*)	4.6956	4.4563	5.0882	5.3687
S (NL)	4.6550	4.4797	5.2324	5.6074
S (ENL*)	4.5705	4.3726	4.7001	5.4699
S+P (NL)	4.5850	4.3519	4.8206	5.1344
S+P (ENL*)	4.5587	4.3098	4.7001	5.1267
T(1) (NL)	4.5514	4.3377	4.8929	5.1570
T(1) (ENL*)	4.5453	4.3199	4.7974	5.1486
T(1.2) (NL)	4.5894	4.3587	4.7617	5.0881
T(1.2) (ENL*)	4.5855	4.3343	4.9615	**5.0788**
$c_{2,2}$ (NL)	4.5511	4.3352	4.8897	5.1529
$c_{2,2}$ (ENL*)	**4.5457**	4.3191	4.7952	5.1444
$c_{2,4}$ (NL)	4.5684	4.3424	4.9180	5.1827
$c_{2,4}$ (ENL*)	4.5626	4.3218	4.8105	5.1728
$c_{4,2}$ (NL)	4.5758	4.3388	4.7618	5.1152
$c_{4,2}$ (ENL*)	4.5741	4.3319	4.6875	5.1059
$c_{4,4}$ (NL)	4.5732	4.3297	4.7664	5.1239
$c_{4,4}$ (ENL*)	4.5675	**4.3116**	**4.6732**	5.1127
$c_{6,2}$ (NL)	4.6164	4.3577	4.7558	5.1437
$c_{6,2}$ (ENL*)	4.6159	4.3576	4.6912	5.1360
$c_{2+2,2}$ (NL)	4.5737	4.3374	4.7632	5.1194
$c_{2+2,2}$ (ENL*)	4.5710	4.3338	4.6877	5.1486

Table 2: Entropies of the pyramids for different four-stages transformations.

Image	lady	lacornou	bureau	carrefour
Original	7.2158	7.5747	6.9054	6.2086
OLP	**5.1529**	**5.1174**	4.1409	5.7714
BJPEG	5.1545	5.5130	**4.1242**	**5.5768**
MEP (NL)	5.0722	5.2459	4.6702	5.8879
MEP (ENL*)	5.0665	5.2307	4.6575	5.8449
RLP (NL)	5.0597	5.1927	4.7439	5.9214
RLP (ENL*)	5.0566	5.1805	4.7276	5.8901
S (NL)	5.2200	5.4434	4.5813	5.8589
S (ENL*)	5.1979	5.2545	4.5181	5.4433
S+P (NL)	4.8816	4.9618	4.3155	5.7444
S+P (ENL*)	**4.7613**	4.8476	4.2230	5.3966
T(1) (NL)	4.8730	4.9669	4.2004	5.6927
T(1) (ENL*)	4.8187	4.9239	4.1662	5.4251
T(1.2) (NL)	4.8567	4.8895	4.2272	5.6968
T(1.2) (ENL*)	4.7876	4.8481	4.2035	5.4537
$c_{2,2}$ (NL)	4.8724	4.9653	4.1976	5.6941
$c_{2,2}$ (ENL*)	4.8177	4.9219	**4.1615**	5.4287
$c_{2,4}$ (NL)	4.8783	4.9910	4.2297	5.6986
$c_{2,4}$ (ENL*)	4.8191	4.9183	4.1915	**5.3512**
$c_{4,2}$ (NL)	4.8364	4.8976	4.2549	5.7144
$c_{4,2}$ (ENL*)	4.7936	4.8601	4.1950	5.4537
$c_{4,4}$ (NL)	4.8253	4.9009	4.2611	5.7054
$c_{4,4}$ (ENL*)	4.7639	**4.8272**	4.1935	5.3562
$c_{6,2}$ (NL)	4.8524	4.9088	4.3199	5.7345
$c_{6,2}$ (ENL*)	4.8250	4.8803	4.2476	5.5062
$c_{2+2,2}$ (NL)	4.8481	4.9024	4.2572	5.6987
$c_{2+2,2}$ (ENL*)	4.8064	4.8622	4.1746	5.4737

Table 3: Entropies of the pyramids for different four-stages transformations.

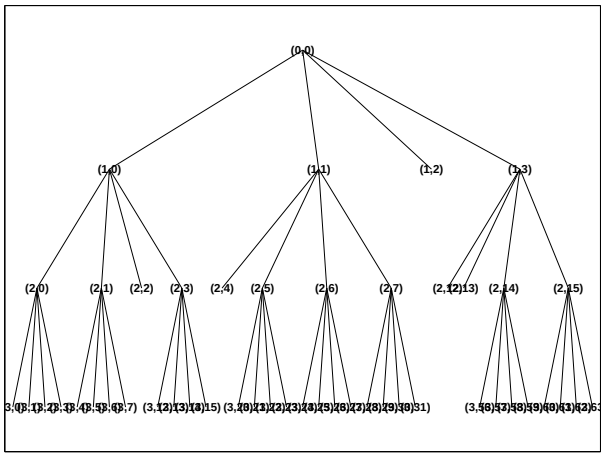

Figure 2: Image "Lena": best tree ($j_m = 3$) associated with ENL* obtained from the NL decomposition $c_{4,4}$.