

# ORIENTED WAVELET TRANSFORM ON A QUINCUNX PYRAMID FOR IMAGE COMPRESSION

Vivien Chappelier, Christine Guillemot

IRISA/Université de Rennes 1, Campus de Beaulieu, 35042 RENNES Cedex FRANCE

Mailto: firstname.lastname@irisa.fr

## ABSTRACT

This paper presents a new image coding technique based on oriented 1D multiscale decompositions on a quincunx sampling grid. A transform obtained by adapting the lifting steps of a 1D wavelet transform along a local orientation is described. A quad-tree structure is used to efficiently code the orientation map, and optimized in the rate-distortion sense. A gain of up to 1.3 dB over the separable wavelet transform is observed for a set of images, with the rate estimated from the stationary entropy of the subbands.

## 1. INTRODUCTION

Wavelet based compression schemes have shown tremendous success due to their capability of approximating signals efficiently with a small number of coefficients. In particular, wavelets are optimal for representing unidimensional signals with a finite number of discontinuities [1]. However, for bidimensional image signals the objects boundaries introduce curve-shaped discontinuities and separable wavelets turn out to be suboptimal. Indeed, in classical wavelet based image coding systems [2], coefficients with high energy are observed around the edges. These coefficients are costly and recent approaches try to overcome this drawback by filtering the image along the contours.

Contourlets [3], for instance, use a fixed directional filterbank to further decompose a multiscale pyramid into oriented subbands. However, this transform is overcomplete, leading to a poor compression performance at high rate. In a different approach, bandelets [4] use an explicit geometric model to describe the edge orientations locally. This model is then used to warp a separable wavelet along the contours. Nevertheless, the bitrate allocation between the geometric model and the coefficients for optimal rate-distortion performance is intricate.

Here, we propose an adaptive transform which consists in applying the lifting steps of a unidimensional wavelet in one of two directions. The direction of filtering is provided locally by an orientation map defined on a quad-tree. This decomposition results in one low-frequency and one high-frequency subband defined on two complementary quincunx

lattices. The transform is applied iteratively leading to a multiresolution decomposition on a quincunx pyramid.

The paper is organized as follows. Section 2 recalls some important results on wavelets and lifting. Section 3 presents the oriented wavelet and the multiresolution decomposition. Section 4 presents the coding of the orientation map and the rate-distortion optimization procedure. Finally, Section 5 evaluates the rate-distortion performance of the oriented wavelet transform compared to the separable wavelet transform for image compression.

## 2. WAVELETS AND LIFTING

Given a signal  $S(t)$ , the dyadic wavelet transform consists in representing  $S$  by its decomposition on a wavelet basis  $(\Psi_{k,n})_{k \in \mathbb{N}, n \in \mathbb{Z}}$ :

$$\Psi_{k,n}(t) = \frac{1}{\sqrt{2^k}} \Psi\left(\frac{t - 2^k n}{2^k}\right)$$
$$S(t) = \sum_{k,n} c_{k,n} \Psi_{k,n}(t)$$

where  $\Psi(t)$  is the *mother wavelet* and  $\Psi_{k,n}(t)$  are scaled and shifted versions of this wavelet. In order to keep both linear phase and perfect reconstruction properties, only biorthogonal wavelets are considered. The transform is implemented using a two-channel FIR filterbank, iterated on the low-pass band.

In [5], Daubechies *et al* prove that the polyphase matrix of such a filterbank can be factorized in a finite number of predict ( $P_i$ ) and update ( $U_i$ ) steps, followed by a final scaling step ( $R$ ). This *lifting* decomposition generalizes the notion of wavelets allowing for the construction of second generation wavelets on sets where the notion of frequency is not defined. Furthermore, the lifting scheme provides an efficient way of implementing the filterbank and guarantees perfect reconstruction by simple inversion of the different lifting steps.

The analysis stage is performed by first splitting the sampling grid in two cosets:

$$C : \begin{cases} S_0[n] := S[2n], \\ S_1[n] := S[2n + 1]. \end{cases}$$

A predict step consists in predicting each sample of  $S_1$  from its neighbors in  $S_0$ :

$$P(\alpha_i) : S_1[n] := S_1[n] + \alpha_i(S_0[n] + S_0[n+1]). \quad (1)$$

Similarly, an update step consists in adjusting the value of each sample in  $S_0$  from its neighbors in  $S_1$ :

$$U(\beta_i) : S_0[n] := S_0[n] + \beta_i(S_1[n] + S_1[n-1]). \quad (2)$$

The predict and update steps are applied successively in alternance. To obtain properly normalized filters, a final scaling step is applied on each coset:

$$R(\chi) : \begin{cases} S_0[n] := \chi S_0[n], \\ S_1[n] := \frac{1}{\chi} S_1[n]. \end{cases}$$

$S_0$  is an approximation of  $S$  and corresponds to a low-pass filtering followed by downsampling, while  $S_1$  represents the prediction error and corresponds to a high-pass filtering followed by downsampling.

The original signal  $S$  is recovered by inverting the lifting steps, that is by applying successively  $R(\frac{1}{\chi})$ ,  $U(-\alpha_{N-1})$ ,  $P(-\beta_{N-1})$ , ...,  $U(-\alpha_0)$ ,  $P(-\beta_0)$  and merging  $S_0$  with  $S_1$ :

$$C^{-1} : \begin{cases} \hat{S}[2n] := S_0[n], \\ \hat{S}[2n+1] := S_1[n]. \end{cases}$$

### 3. ORIENTED WAVELETS ON QUINCUNX GRID

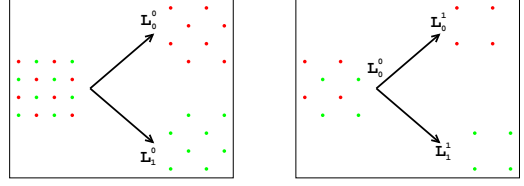
Let us consider a 2D lattice defined from an integer matrix  $M$  as:

$$L(M) = \{m \in \mathbb{Z}^2, m = Mn, n \in \mathbb{Z}^2\}.$$

A multiresolution representation of a 2D signal can be obtained by iterating a 2-channel filterbank with the corresponding downsampling on this 2D lattice. In particular, the square sampling grid  $\mathbb{Z}^2$  can be split in two complementary quincunx grids  $L_0^0 = L(M)$  and  $L_1^0 = L(M) + e$ , where  $M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$  and  $e = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (Fig. 1). This process can be repeated recursively on the  $L_0^0$  grid so as to obtain the following multiresolution structure:

$$\begin{cases} L_0^k = L(M^{k+1}), \\ L_1^k = L(M^{k+1}) + M^k e. \end{cases}$$

The sampling grid is either square or quincunx, for even or odd levels respectively ( $L(M^2) = L(2I) = 2\mathbb{Z}^2$ ). Moreover, a quincunx lattice  $L(M^k)$  ( $k$  odd) can be seen as a square lattice rotated by  $\frac{\pi}{4}$  where the distance between samples is  $2^{\frac{k}{2}}$ . Thus, this recursive partitioning defines an  $l$ -level quincunx sampling pyramid  $(L_1^0, \dots, L_1^{l-1}, L_0^{l-1})$  where the downsampling factor at each scale is  $|\det(M)|^{\frac{1}{2}} = \sqrt{2}$ . This structure has been used previously to define quincunx



**Fig. 1.** Split of the 2D sampling grid in two cosets.

wavelet bases [6]. Unlike the separable case, only one mother wavelet (instead of three) is needed to represent the signal.

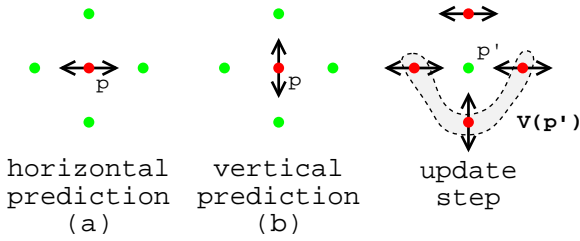
Instead of using quincunx wavelets, the approach here consists in applying a 1D wavelet transform along directions selected adaptively for each wavelet coefficient. These coefficients, corresponding to the prediction errors at the successive levels, are computed on the  $L_1^k$  grids. The approximation images are computed on the  $L_0^k$  grids where the decomposition is iterated. Each point in  $L_1^k$  has 4 neighbors. Therefore, a point  $p \in L_1^k$  can be predicted either from its neighbors in the same line *or* in the same row (Fig. 2). The orientation which minimizes the prediction error is chosen, defining a binary map  $M_k$  on  $L_1^k$ . To compute the wavelet coefficient at location  $p$ , the predict steps of a 1D biorthogonal wavelet are applied in the orientation chosen in  $p$ , as done in a classical 1D lifting implementation. However, the update steps have to be modified. A sample at location  $p' \in L_0^k$  may indeed be used 0 to 4 times to predict its neighbors in  $L_1^k$ , unlike in the 1D case where it is used exactly twice. The factors  $\beta_i^*$  used in the modified update steps are obtained by weighting the original factors  $\beta_i$  of the 1D wavelet given in Eq. 2. Depending on the number  $v$  of neighbors using a sample at location  $p'$  as a predictor, the update factors  $\beta_i^*$  are defined as:

$$\beta_i^* = \begin{cases} \frac{2}{v} \beta_i & \text{if } v \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The update step is thus modified as follows:

$$U(\beta_i) : S_0[n] := S_0[n] + \beta_i^* \sum_{n^* \in V(n)} S_1[n^*],$$

where  $V(n) \subset L_1^k$  is the set of neighbors of  $n$  using a sample at location  $n$  as a predictor. When the direction of prediction is the same for all points in  $L_1^k$ , the decomposition is equivalent to the 1D wavelet applied in that direction on  $L_1^k$ . Perfect reconstruction is guaranteed by inverting each lifting step, though it requires the knowledge of the orientation map  $(M_k)_{k=0..n-1}$ . Moreover, if  $L_1^k$  can be partitioned in a small number of regions where the filtering direction is constant (homogeneous map) and if the 1D wavelet is quasi-orthogonality, then this property is preserved for the oriented wavelet, allowing for easier rate-distortion optimization.

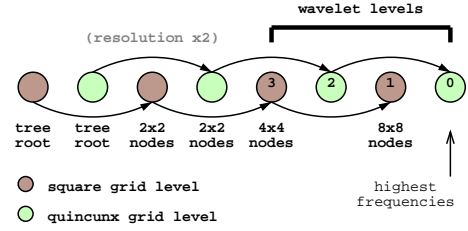


**Fig. 2.** The predict step is applied either in the horizontal (a) or vertical (b) direction. The update step depends on the directions chosen for its four neighboring pixels.

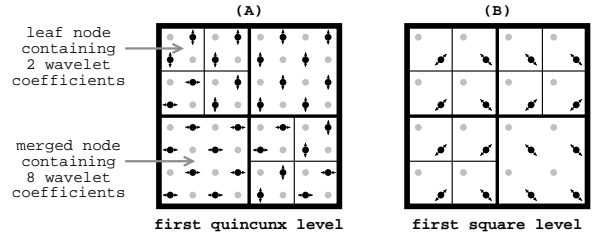
#### 4. CODING AND OPTIMIZATION OF THE ORIENTATION MAP

In order to compact the energy as much as possible in the lower subbands, the orientation map is chosen so as to minimize the prediction error at each level. Without entropy coding, the orientation map would cost slightly less than 1 bpp which is prohibitive. However, this binary information on the filtering direction is not always relevant. Indeed, when the distortion obtained by predicting in either one or the other direction is similar, the choice of the proper orientation does not impact the distortion significantly. This happens mainly in uniform regions, where both predictors are similar, or in textured regions where the pixels are less correlated, hence where the prediction fails. Thus, the orientation information is only important on edges, which concerns a small proportion of the pixels in natural images. It is therefore possible to propagate the orientation information from edges to other regions to reduce the entropy of the map substantially with a negligible impact on the overall distortion. In order to do so, the orientation map is coded using two independent quad-trees, one for horizontal and vertical orientations at even levels and one for diagonal and antidiagonal orientations at odd levels (Fig. 3). Besides, the parent to children relationship captures the strong dependency of the orientations across scales. In the first wavelet decomposition (level 0), the wavelet coefficients are defined on a quincunx lattice and oriented either vertically or horizontally (Fig. 4). For quincunx levels, the two coefficients corresponding to the same leaf in the quad tree share the same orientation.

For sake of simplicity and efficiency, simple linear interpolation is used to compute the predictors during the map optimization. The map is first initialized by the orientation minimizing the prediction error, leading to a very heterogeneous map. The rate  $R(n, m)$  and distortion  $D(n, m)$  associated to each node  $n$  of the tree with orientation  $m$  is computed using the following model based on the low-rate approximation in [4]:



**Fig. 3.** Interlevel quad-tree dependency for a 4-level oriented wavelet decomposition on a 16x16 image.



**Fig. 4.** Wavelet coefficients location for the first (A) and second level (B) of decomposition. The corresponding low-pass coefficients are grayed out. Quad-tree nodes are represented by squares of variable weight with the root in bold.

$$R(n, m) = \begin{cases} \gamma_0 & \text{if } e(n, m)^2 > \sigma \\ 0 & \text{otherwise,} \end{cases}$$

$$D(n, m) = \begin{cases} 0 & \text{if } e(n, m)^2 > \sigma \\ e(n, m)^2 & \text{otherwise,} \end{cases}$$

where  $e(n, m)$  is the prediction error at node  $n$  for orientation  $m$ ,  $\sigma = \frac{4}{3}\lambda\gamma_0$  and  $\gamma_0$  approximates the cost in bits of a coded coefficient. Here we chose  $\gamma_0 = 6$ .

The map is then optimized using a bottom-up approach by computing the lagrangians:

$$J_{merge}(n) = \min_m \left( \sum_{f \text{ child of } n} D(f, m) + \lambda R(f, m) \right),$$

$$J_{split}(n) = \sum_{f \text{ child of } n} D^*(f) + \lambda R^*(f)$$

and using the same orientation for the parent and its children if  $J_{merge} < J_{split} + \Delta$ . The penalty  $\Delta$  represents the additional map cost introduced by splitting the node and is set to  $8L$ , where  $L$  is the quad-tree level ( $L = 0$  for the root). The distortion  $D^*(n)$  and rate  $R^*(n)$  associated to the node are updated according to the decision whether the node is split or not. This process is iterated a few times with  $\lambda$  varying until the target rate is reached.  $D^*(n)$  and  $R^*(n)$  are initialized to  $\min_m D(n, m)$  and  $\min_m R(n, m)$  for the leaves.

## 5. RESULTS

The 1D oriented transform described above is compared to the separable wavelet transform in terms of rate-distortion performance in an image coding system. In both cases, a 9/7 biorthogonal wavelet is used. The orientation map is coded simply by using one bit for the orientation of a node and one bit to indicate if it is a leaf. In all our experiments, the cost of the map was marginal, generally below 0.01 bpp. The rate is estimated from the stationary entropy of the subbands for both the separable and oriented wavelet. The cost of the map is added to this rate in the case of the oriented wavelet.

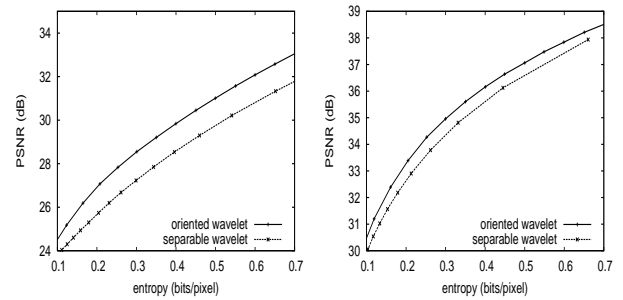
Figure 5 shows the gain obtained by using the proposed transform on the image *barbara* for a rate of 0.3 bpp. Contours are subjectively better preserved and smoother. Figure 6 presents the rate-distortion curves for both the separable and oriented wavelet transform applied on the *barbara* and *lena* images. A gain of more than 1 dB for *barbara* and 0.5dB for *lena* is observed over a wide range of bitrates. At very low rates ( $< 0.2$ bpp), the gain is smaller due to the fixed cost of the orientation map. Table 1 presents the coding performance obtained for a 3-level separable wavelet transform and a 6-level oriented transform on various images. The oriented wavelet transform performs up to 1.3 dB better than the separable wavelet, particularly for images with directional features.

image	lena	barbara	bike	goldhill
separable (dB)	33.98	26.93	23.66	30.51
oriented (dB)	34.62	28.24	24.26	30.50

**Table 1.** Comparison of the separable and oriented wavelet in terms of PSNR at 0.3bpp for a 3-level (resp. 6-level) decomposition.



**Fig. 5.** [left] 3-level separable 9/7 wavelet coding of *barbara* 512x512 at 0.3 bpp, PSNR = 26.93 dB. [right] 6-level oriented 9/7 wavelet coding at 0.3 bits/pixels, PSNR = 28.24 dB. The low pass band has the same resolution in both cases.



**Fig. 6.** Rate-distortion performance for *barbara* (left) and *lena* (right) with a 5-level separable wavelet decomposition and 10-level oriented wavelet decomposition.

## 6. CONCLUSION

In this paper, we have described a new transform for image compression using unidimensional wavelets oriented on a multiresolution quincunx sampling pyramid. The filtering direction is defined by an orientation map optimized using a rate-distortion criterion. Using a simple rate estimation, a large gain in rate-distortion performance is observed over methods based on separable wavelets.

The orientation map introduces a fixed cost which could be reduced at low rates by designing a scalable coding of the map. The map cost could also be reduced using entropy coding. Rate estimation using first order entropy of the coefficients does not take into account the residual dependency captured by state of the art statistical coders. We are currently adapting the EZBC and EBCOT coders to the oriented wavelet to obtain a complete image compression system for a more realistic rate comparison.

## 7. REFERENCES

- [1] M. Vetterli, "Wavelets, approximation and compression," in *IEEE Sig. Proc. Magazine*, Sept. 2001, pp. 59–73.
- [2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," in *IEEE Trans. Image Process.*, Apr. 1992, vol. 1, pp. 205–220.
- [3] M. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *IEEE Proc. Int. Conf. Image Proc.*, 2001.
- [4] E. Le Pennec and S. Mallat, "Sparse geometric image representation with bandelets," in *submitted to IEEE Trans. on Image Process.*, 2003.
- [5] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," in *J. Fourier Anal. Appl.*, 1998, vol. 4(3), pp. 247–269.
- [6] A. Gouze, M. Antonini, and M. Barlaud, "Quincunx lifting scheme for lossy image coding," in *IEEE Proc. Int. Conf. Image Proc.*, Sept. 2000, vol. 1, pp. 665–668.