IMAGE CODING WITH ITERATED CONTOURLET AND WAVELET TRANSFORMS

Vivien Chappelier, Christine Guillemot and Slavica Marinković

IRISA/Université de Rennes 1, Campus de Beaulieu, 35042 RENNES Cedex FRANCE
Mailto: firstname.lastname@irisa.fr

ABSTRACT

This paper presents a new coding technique based on a mixed contourlet and wavelet transform. The redundancy of the transform is controlled by using the contourlet at fine scales and by switching to a separable wavelet transform at coarse scales. The transform is then optimized through an iterative projection process in the transform domain in order to minimize the quantization error in the image domain. A gain of respectively up to 0.5dB and to 1 dB over respectively contourlet and wavelet based coding has been observed for images with directional features.

1. INTRODUCTION

Overcomplete transforms have been used extensively in various signal processing domains such as denoising because of their increased flexibility with respect to critically sampled transforms. However, they do not seem to be a natural choice for compression, as the output space is larger than the input space. Nevertheless, in a non-linear approximation context, a few coefficients of an overcomplete transform can approximate the signal faster than a critically sampled transform due to the increased freedom in the design and choice of the set of coding atoms. Given an arbitrary set of basis function, the Matching pursuit [1] algorithm is a well-known technique to find a sparse representation close to the signal, however it is computationally intensive and sub-optimal.

By introducing structure in the overcomplete atom set, more efficient algorithms using linear transforms can be designed based on the theory of frames, still with increased flexibility over critically sampled transforms. Such designs include overcomplete steerable transforms [2], Gabor transforms or complex wavelet transforms [3], which are all highly redundant.

In the context of image coding, one well-known limitation of the critically sampled transforms such as the discrete wavelet transform (DWT) is the lack of directionality leading to poor performance in the coding of contours. Thus, in order to improve the performance of transform coding, one has to capture the geometrical structure of the image, either explicitly by coding the contours [4], or implicitly by using directional transforms [5]. The contourlet transform [6] provides a multiresolution analysis, with an arbitrary number of directional bands at each level, and forms a tight frame with a small redundancy factor.

By using the theory of projection onto convex sets in [3], Kingsbury et al. propose an iterative algorithm to minimize the distortion introduced by quantization in the overcomplete complex wavelet transform domain. A significant gain over DWT is achieved, however the number of directions is fixed and the redundancy of the transform is high (4:1). As a result, overcomplete directional transforms provide efficient non-linear approximation at lower rate, but are outperformed by critically sampled transforms at higher rate.

Here, we propose an hybrid scheme in which the lowest frequencies are coded using a critically-sampled separable wavelet transform, and the higher frequencies are coded using a slightly redundant contourlet transform. Further minimizing the error introduced in the overcomplete transform by using the iterative projection algorithm on the contourlet transform domain provides significant gains over the DWT with the benefit of a fine-grain directional analysis for a comparable complexity.

The paper is organized as follows. Section 2 presents the algorithm for selection between wavelet or contourlet coding. Section 3 presents the iterative projection algorithm applied to contourlets. Finally Section 4 presents the improvements over DWT and contourlets for non-linear approximation of images.

2. HYBRID WAVELET/CONTOURLET CODING

Introduced in [6], the discrete contourlet transform provides a multiresolution and directional analysis of a 2D signal by first applying a redundant pyramidal multiresolution decomposition followed by a critically sampled directional filterbank applied on each of the undecimated high frequency bands (Fig. 1). At each stage of the pyramidal decomposition, only the low-pass band is downsampled by 2 vertically and horizontally, leading to a redundancy of $(1 + \frac{1}{4}) : 1$ per level of decomposition, i.e. at most $(1 + \frac{1}{3}) : 1$ for the whole transform as the number of decomposition levels tends towards infinity. The low-pass filters used for analysis and

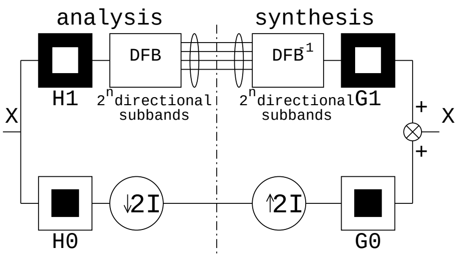

Fig. 1. One stage of the contourlet transform. Black regions indicate the ideal filter frequency response. The high pass band after filtering by H1 is not downsampled.

synthesis are the usual 9/7 tap filters from Antonini et al.[7], and the synthesis is the dual of the analysis frame operator. The directional decomposition is obtained from n iterations of a 2-channel fan filterbank, with prior resampling to get the 2^n oriented subbands (Fig. 2).

Although contourlets are very efficient in approximating the image contours with only a small number of coefficients, their efficiency on natural images with regards to wavelets tend to drop when the number of coefficients increases. Indeed, in the extremal case of lossless coding, the energy of the image is spread in a higher number of coefficients, and wavelet coding performs better. As natural images generally have more energy in the lower frequencies, for a given bitrate there is usually much more wavelet or contourlet coefficients in the coarse scales than in the finer scales. This suggests using a wavelet decomposition to code the coarse scales and a contourlet decomposition for the fine scales. Since the low-pass band of a contourlet transform is exactly the same as with the 9/7 wavelet transform, each level can be decomposed into either 3 wavelet subbands or 2^n directional subbands independently. Moreover, as energy tends to decrease from coarse to fine scales, coding the index of the level at which the transition between contourlet and wavelet occurs is enough and its cost is negligible.

The switching level is determined using a rate-distortion criterion by minimizing the lagrangian $J(\lambda) = D + \lambda R$, where D is the distortion in the transform domain and R is estimated from the absolute entropy of the quantized wavelet or contourlet coefficients in each subband. At low bitrates, $\lambda \approx \frac{3\Delta^2}{4\gamma_0}$ [8], with $\gamma_0 = 7$ and Δ the dead zone quantizer step size. Then the quantizer step size is the same for every subbands. The optimization amounts to searching for Δ and a corresponding switching level leading to the desired rate R. The appropriate switching level for a given Δ is determined by an exhaustive search for the configuration minimizing the distortion among the $L+1$ possibilities, with L the number of decomposition levels.

This new transform provides a directional analysis in the high frequencies with improved performance over wavelets

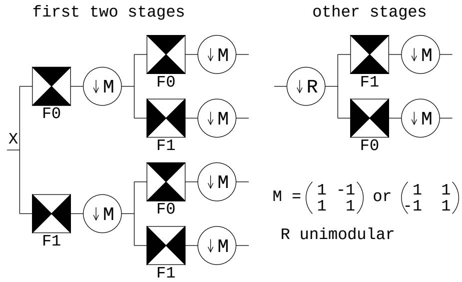

Fig. 2. Directional analysis filterbank. At each stage the signal is critically downsampled by 2 on the quincunx lattice generated by M. $F0, F1$ is a biorthogonal fan filter pair.

while keeping scalability. Figure 3 illustrates the frequency split performed by the transform assuming ideal filters.

3. ITERATIVE CONTOURLET OPTIMIZATION

Overcomplete linear transforms such as the contourlet transform are frame operators. A frame $\{\psi_k\}_k$ is a family of functions in an Hilbert space H such as

$$\forall f \in H, A\left\|f\right\|^2 \leq \sum_k \left|\langle f, \psi_k \rangle\right|^2 \leq B\left\|f\right\|^2$$

with $0 < A \leq B < \infty$. It was shown in [9] that the laplacian pyramid, used for the multiresolution decomposition of the contourlet transform, provides a frame, with $A = B = 1$ if the filters are orthogonal. Hence, the contourlet transform based on the above directional filter bank makes a frame which is *tight* if the filters are orthogonal. There exists a dual frame operator which provides optimal linear reconstruction. For a signal x represented as an N-dimensional vector, and A the MxN matrix of the analysis operator, the dual operator is obtained from the pseudo-inverse $B = (A^T A)^{-1} A^T$ and is shown to minimize the reconstruction MSE among all reconstruction operators in the presence of additive white noise in the transform domain.

In a compression application, the quantization process, expressed as a non-linear function $q(y, \Delta)$ on $y = Ax$,

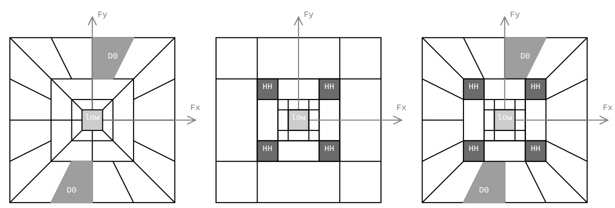

Fig. 3. Frequency partitioning of contourlet, wavelet and hybrid scheme. One directional subband at the 1st level and the HH subband at the 2nd level are highlighted in gray.

introduces a noise d which is neither additive nor uncorrelated. Since the transform is overcomplete, there exists many vectors y that reconstruct the same input signal x. Let $\mathcal{S} = \{y = Ax, x \in \mathbb{R}^N\}$ and $\mathcal{S}^\perp$ its orthogonal complement in $\mathbb{R}^M$. Then

$$\forall y \in \mathbb{R}^M, \exists!(y^\mathcal{S}, y^\perp) \in \mathcal{S} \times \mathcal{S}^\perp, y = y^\mathcal{S} + y^\perp$$

with $y^\mathcal{S} = ABy = P^\mathcal{S}y$ and $y^\perp = (I - AB)y = P^\perp y$. Since the family of basis functions $\{\psi_k\}_k$ is overcomplete, the transform of the error $A(x - \hat{x}) = A(x - B\hat{y}) = A(x - B(y + d)) = A(x - BAx - Bd) = -ABd = -P^\mathcal{S}d = -d^\mathcal{S}$ differs from the distortion d initially introduced by the non-linearity q. Approximating dead-zone quantization by thresholding, which is only valid at low rates, $d^\mathcal{S}$ may also contain non-zero coefficients at locations where $\hat{y}$ is non-zero. Thus subtracting $d^\mathcal{S}$ from $\hat{y}$ tends to increase the energy of kept coefficients while reducing the energy of thresholded coefficients, and $B(\hat{y} - d^\mathcal{S}) = B(y + d^\perp)$ is still equal to x. For the same number of non-zero coefficients, the sparse vector obtained by thresholding $\hat{y} - d^\mathcal{S}$ is closer to y than $\hat{y}$, thereby reducing the reconstruction error.

The iterative projection algorithm designed in [3] for the overcomplete complex wavelet transform is based on this observation and can also be used in the context of contourlets or in the proposed hybrid wavelet/contourlet transform. The input image x is transformed to y_0. Each loop consists in the following steps: y_i is quantized to $\hat{y}_i$ introducing distortion d_i, $\hat{x}_i = B\hat{y}_i$ is reconstructed by pseudo-inverse transform and subtracted from x to obtain the error image e_i. The update coefficients u_i are obtained by transforming e_i and added to $\hat{y}_i$ to form y_{i+1}. The process is iterated a few times until convergence. Noting that $y_i^\mathcal{S} = y_0$ leads to the recursive relation between y_i and y_{i+1} :

$$\begin{aligned} u_i &= A(x - B\hat{y}_i) = y_0 - y_i^\mathcal{S} - d_i^\mathcal{S} \\ y_{i+1} &= \hat{y}_i + u_i = y_i + d_i + y_0 - y_i^\mathcal{S} - d_i^\mathcal{S} \\ &= y_0 + y_i^\perp + d_i^\perp = y_i + d_i^\perp \end{aligned}$$

Using thresholding and a given set of kept coefficients convergence is ensured by using the theory of projection onto convex sets [10][3], thus $d_i^\perp$ tends to zero, meaning the noise is entirely contained in $\mathcal{S}$. Although the theoretical conditions are not met with dead-zone quantization and a varying set of non-zero coefficients, the algorithm still converges when applied to the presented hybrid contourlet/wavelet scheme on natural images.

4. RESULTS

Unless stated otherwise, the results are presented on a 4-level decomposition, using the usual 9/7 biorthogonal filters for multiresolution pyramidal and wavelet filtering. The filters used for the directional decomposition are the biorthogonal FIR filters from Phoong et al. [11], with $N = 16$. The number of directional subbands at each level is 16, which provided the best results. The rate is estimated from the absolute entropy of the quantized subbands. The same quantizer step size is used in every subbands and the distribution is assumed stationnary and estimated from the subbands histograms. Since the low-pass band is the same for every considered transform, it is not quantized nor taken into account in the rate estimate.

image	bpp	wavelet	contourlet	hybrid	iterated
barbara	0.200	26.62	26.83	27.34	27.57
bike	0.224	23.31	23.14	23.31	23.87
zoneplate*	0.232	17.31	17.43	18.08	18.21

Table 1. PSNR (dB) values obtained with the different transforms applied on two 512x512 natural images and a synthetic image (*) made of concentric circles.

Table 1 shows the performance improvement over wavelet and contourlet coding for a few images. These images contain directional features in the high frequencies that are well captured by the directional filterbank at finer scales. The lower frequencies are coded efficiently with wavelets. This results in a 0.2 to 0.5 dB gain over standard contourlet transform based coding. The iterative optimization procedure provides an additional 0.2dB improvement. For images without directional features, the hybrid transform falls back to wavelet transform thanks to the rate-distortion optimization process, guaranteeing performance no worse than wavelet coding for any image.

Figure 5 and 6 illustrate the gain of the presented transform on the barbara image at varying rates. A gain of 0.5dB to 1dB over wavelet coding is observed over a wide range of bitrate for the hybrid iterated transform. Most of the gain over the non-iterated transform is obtained after only the first iteration.

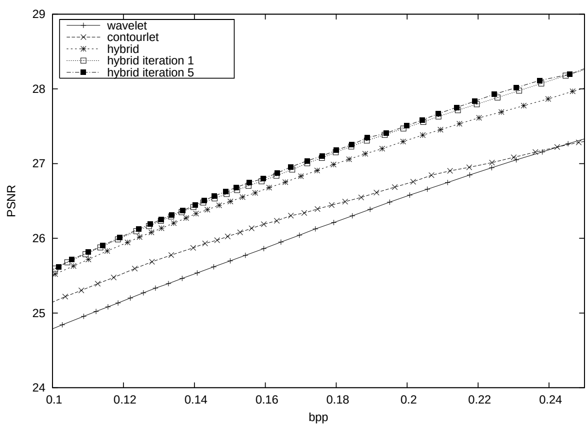

Fig. 4. Rate-distortion performance of the wavelet, contourlet, hybrid and iterated hybrid transforms on barbara.

Fig. 5. 4-level 9/7 wavelet coding of barbara at 0.2 bpp, PSNR = 26.62 dB.

Fig. 6. 4-level iterated hybrid wavelet/contourlet coding of barbara at 0.2bpp, PSNR = 27.56 dB.

5. CONCLUSION

A new transform using overcomplete directional filtering to capture high frequency directional features and wavelets to efficiently code the lower frequencies has been presented. The level at which the transition between directional analysis and wavelet analysis occurs is determined using a rate-distortion criterion. An optimization procedure based on iterative projection of the quantization error is applied to this hybrid transform. An overall gain of 0.5 to 1 dB in PSNR over wavelet coding is observed for images with directional features. Further improvements could include optimizing the number of directional subbands during the rate-distorsion optimization procedure and using advanced entropy coders such as EBCOT.

6. REFERENCES

[1] S.G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," in *IEEE Trans. Signal Processing*, Dec. 1993, vol. 41, pp. 3397–3415.

[2] B. Beferull-Lozano and A. Ortega, "Coding techniques for oversampled steerable transforms," in *Proc. Int. Asilomar Conf. on Signals, Systems and Computers*, Oct. 1999.

[3] Nick Kingsbury and Tanya Reeves, "Iterative image coding with overcomplete complex wavelet transforms," in *SPIE Vis. Comm. Image Proc.*, July 2003, vol. 5150, pp. 1253–1264.

[4] E. LePennec and S. Mallat, "Geometrical image compression with bandelets," in *SPIE Vis. Comm. Image Proc.*, July 2003, vol. 5150, pp. 1273–1286.

[5] E.J. Candes and D.L.Donoho, *Curvelets – a surprisingly effective nonadaptive representation for objects with edges*, pp. 1–10, Vanderbilt University Press, Nashville, TN, 1999.

[6] M. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *IEEE Proc. Int. Conf. Image Proc.*, 2001.

[7] 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.

[8] Erwan LePennec and Stephane Mallat, "Sparse geometric image representation with bandelets," in *submitted to IEEE Trans. on Image Process.*, 2003.

[9] M. N. Do and M. Vetterli, "Framing pyramids," in *IEEE Transactions on Signal Processing*, Sept. 2003.

[10] Y.Yang and N.P. Galatsanos, "Removal of compression artifact using projections onto convex sets and line process modelling," in *IEEE Trans. Image Proc.*, Oct. 1997, vol. 6, pp. 1345–1357.

[11] S. M. Phoong, C.W. Kim, P.P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," in *IEEE Trans. SP*, Mar. 1995, vol. 43(3), pp. 649–665.