

HIDDEN MARKOV TREE IMAGE DENOISING WITH REDUNDANT LAPPED TRANSFORMS

Laurent Duval

Institut Français du Pétrole
Technology Department
92852 Rueil-Malmaison Cedex, France
laurent.duval@ifp.fr

Truong Q. Nguyen

Department of ECE
University of California, San Diego
La Jolla, CA 92093-0407
nguyent@ece.ucsd.edu

ABSTRACT

Hidden Markov trees (HMT) wavelet models have demonstrated superior performance in image filtering, by their ability to capture features across scales. Recently, we proposed to extend the HMT framework to the Lapped Transform domain, where Lapped Transforms (LT) are M -channel linear phase filter banks. When the number of channels is a power of 2, the block partition provided by LT is remapped to an octave-like representation, where an HMT is able to model the statistical dependencies between intra- and inter-band coefficients. Due to better energy compaction and reduced aliasing properties, LT outperforms discrete wavelet transforms at moderate noise levels, both subjectively and objectively. However, critically-decimated LT suffers from a lack of shift-invariance, resulting in a degraded performance. In this paper, we study the improvement of HMT modeling in the LT domain (HMT-LT), combined with a redundant decomposition, in order to increase its performance for image denoising.

1. INTRODUCTION

Sparse representation is a key property in many signal processing algorithms. The discrete wavelet transform (DWT) provides such representations for a lot of real-world signals. As a consequence, numerous DWT-based algorithms form the basis for efficient signal and image statistical analysis, where asymptotically optimal performance is achieved by wavelet-domain thresholding, in the case of Gaussian additive noise [1]. The key to noise filtering is to study signals in domains where statistics of the clean signal and the noise are modeled more efficiently, via appropriate transformations. Wavelet decompositions exhibit two heuristic properties often termed "clustering" and "persistence": feature-related wavelet coefficients (edges or singularities) tend to cluster locally in a subband and to persist across scales, through the wavelet tree. Recently, algorithms adopted tree-adapted subband-dependent shrinkage [2]. Also, sophisticated models of the joint statistics may be useful for capturing key-features in real-world images. A recent approach relies on Markov random fields. We refer to [3] for an rich overview of their use in signal and image processing. Based on the Hidden Markov Tree framework developed in [4], H. Choi *et al.* have proposed efficient image denoising [5].

In [6], we proposed to extend the use of hidden Markov models to a lapped transform (HMT-LT) domain. The use of Lapped Transforms was motivated by their superior energy compaction properties [7] as well as their robustness to oversmoothing. For

our application, the LT coefficients are rearranged into a pseudo-octave representation, bearing the same clustering and persistence properties as in the wavelet representation. In the following, we first briefly review some properties of the Lapped Transforms and the dyadic re-mapping of the transformed coefficients. In Section 3 basic principles behind Hidden Markov Tree modeling are briefly reviewed. We then explain the association of HMT denoising in the Lapped Transform domain and redundant decomposition. Results are given in Section 5, in association with heuristics (Section 6) proposed for computational cost reduction, based on randomized averaging over shifts.

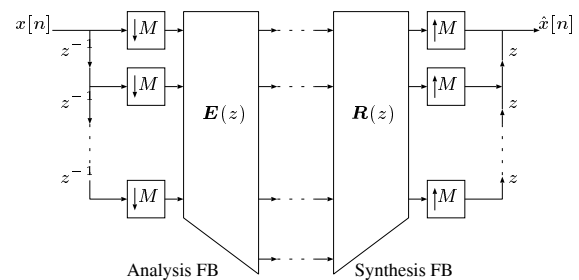


Fig. 1. Block diagram of the polyphase matrices $\mathbf{E}(z)$ and $\mathbf{R}(z)$ of a processing system based on a M -band critically sampled filter bank.

2. LAPPED TRANSFORMS AND OCTAVE REPRESENTATION

The Lapped Orthogonal Transform (LOT, [8]) has been developed to overcome the annoying blocking effects of the DCT. Lapped transforms are linear phase paraunitary filter banks (FB). A block diagram of the analysis and synthesis FB pair is given in Figure 1. The analysis and the synthesis M -band FB polyphase matrices ($\mathbf{E}(z)$ and $\mathbf{R}(z)$ respectively) provide perfect reconstruction with zero delay if and only if:

$$\mathbf{R}(z)\mathbf{E}(z) = \mathbf{I}_M,$$

where \mathbf{I}_M is the identity matrix [9, p. 304]. LT may be parameterized through efficient lattice structures for cost-driven optimiza-

tion. We refer to [8, 9, 7] for a comprehensive overview on Lapped Transforms.

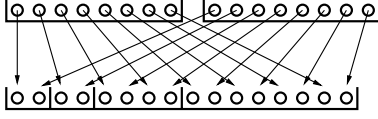


Fig. 2. Dyadic rearrangement of 1D LT coefficients: (Top) block-transform with uniform frequency partition, (Bottom) octave-like representation.

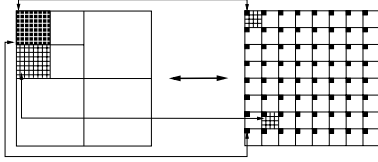


Fig. 3. Dyadic equivalence between octave and block transforms: (Left) two-level octave-like representation, (Right) four-channel block-transform with uniform frequency partition.

Lapped Transforms project signals onto M equally spaced frequency bands, in contrast to the octave-band wavelet representation. When M is a power of 2, the transformed coefficients bear an octave-like grouping, with $J = \log_2 M$ decomposition levels. It follows from the trivial identity $2^J = 1 + \sum_{j=1}^J 2^{j-1}$. The DC component (the "1" term) is assigned to the lower scale subband. Then, from low to high frequencies, the j th subband is formed respectively from the next group of 2^{j-1} coefficients. The $J + 1$ groups are then associated with respect to the block position in the signal. Figure 2 represents the dyadic rearrangement from two consecutive blocks of $8 = 2^3$ coefficients into a three-level decomposition.

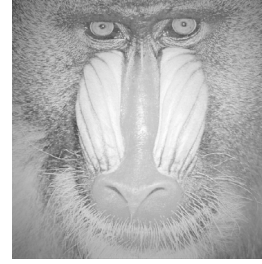
The re-mapping from a four channel block transform to a two level dyadic transform is depicted in Fig. 3. Arrows between coefficients (little black or white squares) link reciprocal locations of coefficients in the dyadic and the block grouping scheme. Once wavelet and LT coefficients share similar grouping, the same denoising procedure may be applied to both domains.

The re-mapping from an eight-channel DCT-II block transform into a three-level pseudo-octave is illustrated on "Mandrill" in Fig. 4. Figure 4-b is made of 8×8 subblocks. Figure 4-c represents its pseudo-octave representation. The coefficient magnitude has been scaled by an exponential factor for visualization.

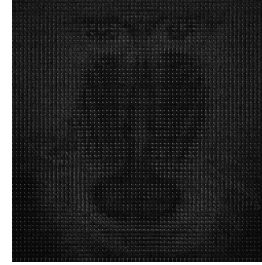
3. HIDDEN MARKOV TREE MODEL FOR OCTAVE DECOMPOSITION

Wavelet coefficients of real-world images generally possess a non-Gaussian distribution, with a lot of small coefficients and few large ones, due to the sparse properties of wavelet decompositions.

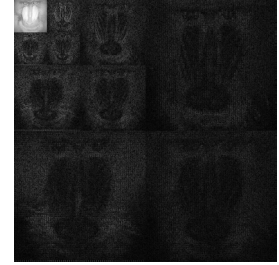
The Hidden Markov Tree (HMT) model developed by Crouse *et al.* [4] is often described as a quad-tree structured probabilistic graph that captures the statistical properties of the wavelet transforms of images. It is based on a hidden Markov model (HMM) which exploits hidden states. The non-Gaussian behavior of the coefficients is modeled as a Gaussian mixture with two components. The hidden states of the HMM are the large (L) or small



(a)



(b)



(c)

Fig. 4. (a) Mandrill image (b) 8-channel block decomposition (c) 8-channel octave decomposition.

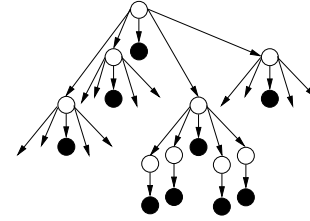


Fig. 5. Diagram of a Hidden Markov Tree in a quad-tree. White dots represent hidden states with arrows as parent-child dependencies, black dots the wavelet coefficients whose conditional distribution depends on the nature of the hidden state.

(S) nature of the coefficients. It determines the conditional distribution of the associated coefficient. Since the coefficient nature tends to propagate across scales (see [10, 5]), the Hidden Markov Tree materializes the cross-scale (parent-child) link between hidden states. A template HMT is depicted in Figure 5. The parameters of the HMT model are estimated using an Expectation Maximization (EM) algorithm. We refer to [4, 5] for details on the implementation of Hidden Markov Trees.

4. REDUNDANT LAPPED TRANSFORMS

For compression purposes, non-redundant transforms are often considered. In other applications such as detection, segmentation or filtering, redundancy is generally recognized as a substantial improvement. In the particular framework of filter banks, the classical critically decimated construction generally induces a sensitivity to integer delays in signals.

Let S be the "unit" shift operator defined on a discrete signal x by $S(x)_j = x_{j+1}$. Due to subsampling, a L -level discrete wavelet transform generally produces different coefficients for the signal x

and its shifted versions $S^k x$, unless k is a multiple of 2^L . Traditional shrinkage thus yields different estimates of the clean signal \hat{x} , depending on the shift k , which may not be realistic in many applications. Moreover, shift-variance is also associated with annoying ringing artifacts in the vicinity of discontinuities in the signal.

Shift-invariant or stationary transforms have been used to "fill in the gaps" (cf. [11]) generated by the subsampling operators. The representation of the original signal thus becomes redundant. One of the most popular implementation consists in applying a denoising scheme D_λ (for instance a thresholding operator) to a range of meaningful signal shifts $k \in I$, and to obtain several denoised estimates $DS^k x$ of x by shifting them back $\hat{x}_k = S^{-k} DS^k x$. A new denoised estimate $\hat{x}_{cs} = E_{k \in I}(\hat{x}_k)$ is then obtained by this "cycle-spinning" procedure; the mean on all shifts is often used [12]. An similar method has also been used for JPEG deblocking [13], and further extended to other collections of transformations, for instance including translations and rotations for images [14]. For M -channel filter banks, it is sufficient to use M shifts to achieve shift-invariance for 1-D signals. In this work, we first used $M \times M$ shifts for 2-D images (see Section 5). The huge computational burden may be reduced by practical considerations (see Section 6).

5. APPLICATION: IMAGE DENOISING

We have performed simulations on five images with different characteristics, at five noise levels. Table 1 compares the performance of the HMT on both shift-variant and shift-invariant filtering via Lapped Transforms.

We have reported in Table 1 the PSNR of the noisy image and the result of "direct" HMT-LT denoising on the first two lines. Direct denoising is performed as described in [6], where Lapped Transform improvements over wavelets is also discussed. But since shift-variant filtering naturally depends on shifts, we also have reported the minimum and the maximum PSNR obtained from filtering of the $M \times M$ possible shifted versions of the original image. They are denoted by Min_{cs} and Max_{cs} . This issue is important, as we can see from the "Boat" image, which exhibit an extreme behavior. With a 17.7 dB noise, the PSNR after shift-variant filtering may vary from 25.8 to 27.0 dB, simply by shifting and reverse shifting after filtering. Such a dramatic change ought to be mentioned for a fair comparison between denoising methods.

The "redundant" performance is obtained by averaging over all the possible shifts. Other more robust estimates could have been used. We have for instance tested a median estimator. Since its results differed in general from the average estimator by less than 0.1 dB, it has not been reported here.

In a majority of the images and noise levels tested here, redundant HMT-LT denoising outperforms direct denoising, with maximum of 1.1 dB improvement for the "Boat" image at low noise level. When the noise level increases, the gain is reduced: it is about 0.5 dB for noise level between 15 to 20 dB, and falls to 0.1-0.2 dB at 12 dB noise (last column).

If we now focus on Min_{cs} and Max_{cs} , we remark that generally shift-variance increases with the level of noise: the difference between Min_{cs} and Max_{cs} increases as the PSNR of the noisy image decreases. As a result, especially at high noise levels, it becomes desirable to find the best shift, i.e. the one that achieves the best PSNR after denoising. As we can see from Table 1, the redundant HMT-LT performs equally or better than the best shift HMT-LT

(except for the "Boat" image with 12.0 dB noise). Interestingly for "Mandrill" at low noise (30.5 dB), both Direct and best shift denoising actually degrade the final noise level. This may be due to the specificity of the animal hair regions, which are difficult to distinguish from noise. In this case, redundant HMT-LT denoising still slightly improves the PSNR by 0.2 dB.

The overall objective improvement of the redundant HMT-LT is illustrated subjectively in Fig. 6, where slightly better texture preservation is observed.

<i>Barbara</i>					
Noisy	30.4	24.4	20.9	17.7	12.0
Direct	33.3	29.1	26.6	24.2	20.0
Min _{cs}	33.3	28.9	26.4	23.9	19.5
Max _{cs}	33.4	29.2	26.6	24.3	20.1
Redundant	34.1	29.9	27.3	24.8	20.2
<i>Boat</i>					
Noisy	30.5	24.4	20.9	17.7	12.0
Direct	36.5	32.4	29.0	26.7	21.4
Min _{cs}	36.3	31.8	28.7	25.8	20.9
Max _{cs}	36.8	32.5	29.6	27.0	21.7
Redundant	37.6	32.9	29.8	27.1	21.6
<i>Goldhill</i>					
Noisy	30.4	24.4	20.9	17.7	12.1
Direct	34.0	30.2	27.9	25.5	21.4
Min _{cs}	34.0	29.9	27.5	25.0	20.6
Max _{cs}	34.1	30.3	27.9	25.7	21.5
Redundant	34.6	30.5	28.1	25.8	21.5
<i>Lena</i>					
Noisy	30.5	24.4	20.9	17.8	12.1
Direct	33.8	29.6	27.2	24.9	20.7
Min _{cs}	33.7	29.5	27.0	24.5	20.1
Max _{cs}	33.8	29.7	27.4	25.1	20.8
Redundant	34.7	30.3	27.8	25.3	20.9
<i>Mandrill</i>					
Noisy	30.5	24.5	20.9	17.7	12.1
Direct	30.2	26.6	24.4	22.3	18.9
Min _{cs}	29.9	26.5	24.4	22.2	18.6
Max _{cs}	30.3	26.6	24.4	22.4	19.1
Redundant	30.7	27.0	24.8	22.5	19.1

Table 1. PSNRs (in dB) comparison between non-redundant and redundant HMT-LT for several images.

6. COMPLEXITY REDUCTION

The proposed algorithm only requires to run the denoising algorithm on shifted versions of a single image. Since we want to improve upon the best shift HMT-LT denoising, it seems natural to perform the average estimator over all the $M \times M$ possible shifts. The number of channels M is here a power of two, typically 8 or 16. The number of possible shifts represents a significant computational complexity. The complexity can be greatly reduced by considering the same modeling HMT for all shifts, reducing the EM training cost for the model parameters.

The number of transformations may also be reduced. In practice, not all the possible shifts are required, as long as one still

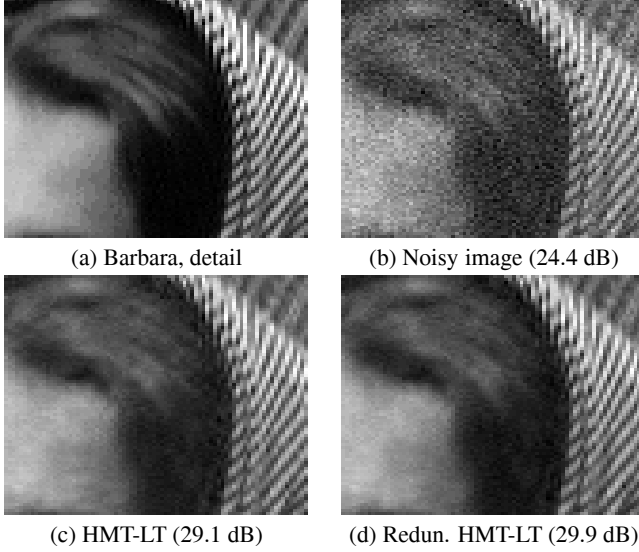


Fig. 6. (a) A portion of the Barbara image (b) with additive noise (24.4 dB), (c) after direct HMT-LT denoising (29.1 dB), (d) after redundant HMT-LT denoising (29.9 dB).

obtains a robust estimate over a subset of shifts. This idea has already been issued in several works, for instance [12] for denoising or [13] in the context of JPEG deblocking, with some shifts chosen *a priori*.

In this work, we used a heuristic based on a randomized strategy proposed in [14]. We have randomly picked k shifts ($k \leq 63$) which are then averaged. Figure 7-a displays three realizations (marked by \times signs) of averaging k randomly picked shifts, k varying from 1 to 63. The black solid line plots the average PSNR over a hundred realization. The horizontal straight lines represent, from bottom to top, the quantities Min_{cs} , Max_{cs} and the PSNR of the proposed algorithm. Few k randomly picked shifts, around 10, suffice to gain over Max_{cs} and reach a PSNR close to that of the redundant HMT-LT. Figure 7-b shows the variation of the PSNR standard deviation between realizations with the number of shifts. The solid line represents the decay we could expect from the averaging of k realizations of a deterministic image corrupted with a random noise.

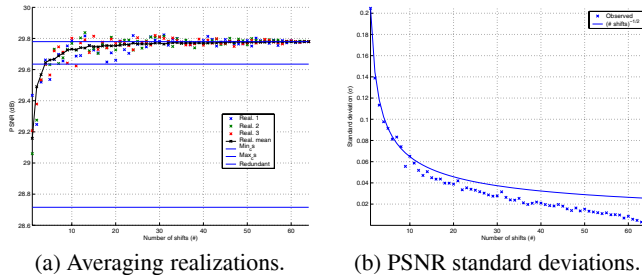


Fig. 7. (a) Three realizations of averaging over a varying number of shifts, from 1 to 63 (colored \times), and comparison with Min_{cs} , Max_{cs} and redundant HMT-LT, for "Boat" at 20.9 dB (b) PSNR standard deviations for hundred randomized shift averaging realizations.

7. CONCLUSIONS

We propose an image denoising algorithm based on a Hidden Markov Tree model applied in the Lapped Transform domain, combined with a redundant decomposition. It is able to outperform the non-redundant HMT-LT algorithm proposed in [6], as well as its "best shift" version. The overall computational cost may be reduced by averaging over a randomized selection of image shifts.

8. REFERENCES

- [1] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. on Inform. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [2] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. on Image Proc.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
- [3] A. Willsky, "Multiresolution Markov models for signal and image processing," *Proc. IEEE*, vol. 90, no. 8, pp. 1396–1458, Aug. 2002.
- [4] M. Crouse, R. Nowak, and R. Baraniuk, "Wavelet-based signal processing using hidden Markov models," *IEEE Trans. on Signal Proc.*, vol. 46, no. 4, pp. 886–902, April 1998.
- [5] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," in *Proc. SPIE Technical Conference on Mathematical Modeling, Bayesian Estimation, and Inverse Problems*, 1999, pp. 31–44.
- [6] L. Duval and T. Q. Nguyen, "Lapped transform domain using Hidden Markov trees," in *Int. Conf. on Image Processing*, 2003.
- [7] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," *IEEE Trans. on Signal Proc.*, vol. 48, pp. 133–147, Jan. 2000.
- [8] Henrique S. Malvar, *Signal Processing with Lapped Transforms*, Artech House, 1992.
- [9] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.
- [10] Jerome M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Proc.*, vol. 41, pp. 3445–3462, Dec. 1993.
- [11] G. P. Nason and B. W. Silverman, *The Stationary wavelet transform and some statistical applications*, pp. 281–300, In Antoniadis and Oppenheim [15], 1995.
- [12] R. R. Coifman and D. L. Donoho, *Translation-invariant Denoising*, pp. 125–150, In Antoniadis and Oppenheim [15], 1995.
- [13] Aria Nosratinia, "Denoising of JPEG images by re-applications of JPEG," *Journal of VLSI Signal Processing*, vol. 27, pp. 69–79, 2001.
- [14] P. Moulin and J. Liu, "On the risk of transformation-averaged wavelet estimators," in *Wavelets and statistics*, Sep. 2003, p. 15, Abs.
- [15] A. Antoniadis and G. Oppenheim, Eds., *Wavelets and Statistics*, Springer-Verlag, 1995.