BRIDGING SCALE-SPACE TO MULTISCALE FRAME ANALYSES

Yufang Bao and Hamid Krim

ECE Dept., NCSU,
Raleigh, NC 27695-7914,
yfbao@eos.ncsu.edu, ahk@eos.ncsu.edu

ABSTRACT

We address a well known problem of nonlinear image diffusion techniques, namely the loss of texture information. We do so by first determining that it is due to unaccounted correlation structure in the image which we subsequently mitigate by proposing a wavelet frame-based technique. This, by the same token establishes a theoretical bridge between the scale space methodology and the multiscale analysis approach. We provide examples to illustrate the effectiveness of the proposed approach.

1. INTRODUCTION

Several developments in image nonlinear filtering have followed as a result of adopting Witkin's proposed equivalence of a heat Partial Differential Equation (PDE)-based evolution of a process and its smoothing by a Gaussian kernel [1]. Such an evolution leads to the so-called linear scale space of a process, parameterized by scale t,

$$\frac{\partial U(t,x)}{\partial t} = \Delta U(t,x), \qquad (1)$$

with x as a spatial vector coordinates. On the other hand Mallat [2] proposed a systematic multiscale analysis framework using Gaussian wavelets as well as orthonormal wavelet basis. A wavelet ability to focus on and localize salient features of a signal, together with its efficient numerical implementation, have gained multiscale analysis a great popularity. While the latter provided a rigorous and flexible framework, with a solid mathematical foundation and a wealth of novel ideas in denoising, segmentation [3, 4, 5], etc., its analytical tractability has always imposed stringent and even unrealistic statistical assumptions. Specifically, the implicit/explicit assumption of independence of the wavelet coefficients, for example, has been key in several theoretical developments, and has also been limiting or unrealistic for some signal classes and down right incorrect for images. The continuous scale approach (or redundant representation), on the other hand, naturally preserves intra-scale correlation information and is amenable to including inter-scale information, hence giving one reasons to believe that such a strategy might be a more adapted and viable approach. Such an approach which is necessarily nonlinear in nature, was first proposed by Perona and Malik in their landmark paper [6] where they aimed at preserving important sharp features of an image, such as edges. The novelty of this approach together with its very promising results triggered a tremendous research activity in computer vision and applied mathematics (see [7] for a good review of the literature) and opened up more research avenues. A number of very good and recent papers have provided inspiring variational interpretations of various nonlinear smoothing techniques [7, 8]. The predominantly deterministic and more recently stochastic analyses [9], share a common limitation of eventually adversely affecting any texture information present in an image. This as further elaborated below, is inherent to the first order Markov property assumed for the image.

This was in part supported by an AFOSR grant F49620-98-1-0190 and by ONR-MURI grant JHU-72298-S2.

In this paper, we address this problem and show that using wavelet frames with wavelets of higher order than Haar is tantamount to accounting for longer term correlation and leads to a good preservation of texture while removing noise and making the image more amenable to other processing. In the next section, we give a brief background review and formulate the problem. In Section 3, we develop a connection between linear filtering using a heat operator and smoothing by Haar frames. In Section 4, we propose a more general frame-based smoothing approach using higher order wavelets such as Daubechies' or others, and provide some substantiating examples in Section 5.

2. BACKGROUND AND PROBLEM STATEMENT

As noted above the Perona-Malik equation still enjoys a great deal of popularity for achieving a selective nonlinear filtering, compatible with the desired objective of image filtering, namely that homogeneous areas be maximally smoothed while edge contours be maximally preserved (or equivalently minimally smoothed). It is expressed as

$$\frac{\partial U(t,x)}{\partial t} = \operatorname{div}\left(g\left(|\nabla(U(t,x))|\right)\nabla U(t,x)\right), \qquad (2)$$

Where $g(v)$ can be chose as $g(v) = e^{-\frac{v^2}{K^2}}$, K determines the rate of decay and thus the extent of smoothing of $U(t,x)$ for a given gradient size. Many other techniques have been proposed and each addressing different aspects of the limitations of the above equation. To the best of our knowledge, none of these techniques [7, 9] resolves the problem of texture loss alluded to in the introduction. The fact that the gradient-based (or Markovian property) selective filtering of an image is a main contributor may easily be observed by the convergence of the data to staircase functions[10] as can be seen in Fig. 1.

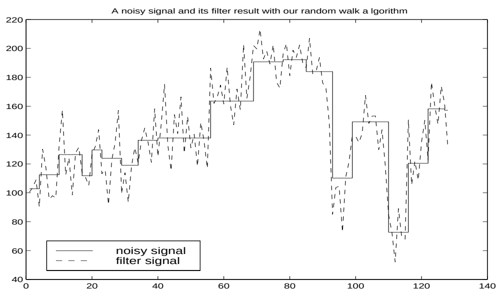

Figure 1: A profile of noisy Lenna image and filtered result with Random walk algorithm.

3. BRIDGING MULTISCALE ANALYSIS AND SCALE-SPACE ANALYSIS

3.1. Frame Representation and Reconstruction

As is well known, an image may be well represented in a Haar wavelet frame by obviating the dyadic down-sampling step of the usual orthonormal representation (or the redundant representation). While the latter representation yields a perfect reconstruction, it experiences a visually noticeable loss of information in the course of a smoothing transformation for noise removal. This is particularly evident when a nonlinear smoothing is locally applied and where a coefficient has a significant impact on the visual outcome. By opting for a redundant representation, we in a sense counterbalance this singular effect. Our goal in this section is to then demonstrate the intrinsic interplay between the smoothing of an image by a linear heat equation and a specific transformation of its frame coefficients. Subsequently using judiciously selected wavelets in the nonlinear filtering, we take advantage of the selected wavelet support size to capture a longer term correlation structure which is crucial to better preserving some features such as texture.

Towards that end, let $\phi(x)$ be a scaling function with a compact support such that$\{\phi(x-n)\}$ is an orthonormal basis of V_0, the space of observations,and $\phi(x)$ with Fourier transform $\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}\hat{h}(\frac{\omega}{2})\hat{\phi}(\frac{\omega}{2})$. $\psi(x)$ be a function with a Fourier transform $\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}(\frac{\omega}{2})\hat{\phi}(\frac{\omega}{2})$ i.e., a so-called mother wavelet. As is well known [11], we can write

$$\frac{1}{\sqrt{2}}\phi(\frac{x}{2}) = \sum_{-\infty}^{+\infty} h(n)\phi(x-n); \frac{1}{\sqrt{2}}\psi(\frac{x}{2}) = \sum_{-\infty}^{+\infty} g(n)\phi(x-n),$$

where $\{h(n)\} and \{g(n)\}$ with Fourier transform$\hat{h}(\omega), \hat{g}(\omega)$ that satisfies $\hat{g}(\omega) = e^{-i\omega}\hat{h}^*(\omega+\pi)$.

Define

$$\begin{cases} \phi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\phi(\frac{x-2^{j-1}n}{2^j}) \\ \psi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\psi(\frac{x-2^{j-1}n}{2^j}) \end{cases}$$

with $\{\phi_{j,n}(x)\}$ and $\{\psi_{j,n}(x)\}$ as tight frames of V_j and W_j (the so-called approximation and detail subspaces). It follows that $\{\phi_{j,2n}(x)\}$ and $\{\phi_{j,2n+1}(x)\}$ as well as $\{\psi_{j,2n}(x)\}$ and $\{\psi_{j,2n+1}(x)\}$ are respectively orthonormal bases of V_j and W_j. As noted above, our goal is to show that a linear heat operation on a process/image may be implemented using the latter's frame coefficients and the following will consequently result.

Proposition 1. *The discrete second difference numerical implementation of a Laplacian operator in a linear heat equation, is equivalent to the highest detail of a wavelet packet frame decomposition of a signal, with a Haar mother wavelet function.*

Prior to proving this proposition, we have to establish the following two lemma.

Lemma 1. *Given a function $f(x)$ together with its frame representation coefficients,*

$$a_j(n) =< f, \phi_{j,n} > \text{ and } d_j(n) =< f, \psi_{j,n} >,$$

and hence the scaling coefficients $\bar{a}_j(n) = a_j(2n)$, in orthonormal bases $\{\phi_{j,2n}(x)\}$ and $\{\phi_{j,2n+1}(x)\}$, we have

$$\begin{aligned} a_{j+1}(p) &= \sum_{-\infty}^{+\infty} h(n-p)\bar{a}_j(n) \text{ and} \\ d_{j+1}(p) &= \sum_{-\infty}^{+\infty} g(n-p)\bar{a}_j(n), \end{aligned} \quad (3)$$

with either of the following two reconstructions of $a_j(n)$

$$\bar{a}_j(p) = \sum_{-\infty}^{+\infty} a_{j+1}(2n)h(2n-p) + \sum_{-\infty}^{+\infty} d_{j+1}(2n)g(2n-p),$$

or

$$\bar{a}_j(p) = \sum_{-\infty}^{+\infty} a_{j+1}(2n+1)h(2n+1-p) + \sum_{-\infty}^{+\infty} d_{j+1}(2n+1)g(2n+1-p)$$

Proof: The proof is immediate by noting that

$$\begin{aligned} &< \phi_{j,p}(u), \phi_{j,2n}(u) >= \\ &\int \frac{1}{\sqrt{2^{j+1}}}\phi(\frac{x-2^j p}{2^{j+1}})\frac{1}{\sqrt{2^j}}\phi(\frac{x-2^j n}{2^j})dx = \\ &\int \frac{1}{\sqrt{2}}\phi(\frac{u-n}{2})\phi(u-n)du \\ &= h(n-p), \end{aligned} \quad (4)$$

and similarly $< \psi_{j+1,p}, \phi_{j,2n} >= g(n-p)$. ■

We now look at the frame representation, and bearing in mind that details of a signal include some signal and noise and that the high frequency portion is mostly noise, which would ideally be removed. Towards that end, we decompose the detail space using a wavelet packet frame as spelled out in the following lemma,

Lemma 2. *The detail space W_j can be expressed as a direct sum of two subspaces $W_j = V_{j,L} \bigoplus W_{j,L}$. Defining*

$$\begin{aligned}\psi^a_{j,p}(x) = & \sum_{-\infty}^{+\infty} h(n-p)\psi_{j,n}(x) \text{ and} \\ \psi^d_{j,p}(x) = & \sum_{-\infty}^{+\infty} g(n-p)\psi_{j,n}(x), \end{aligned} \quad (5)$$

then $\{\psi^a_{j,p}(x)\}$ and $\{\psi^d_{j,p}(x)\}$ are respectively frames of $V_{j,L}, W_{j,L}$ with $(\{\psi^a_{j,2n}(x)\}, \{\psi^d_{j,2n}(x)\})$ and $(\{\psi^a_{j,2n+1}(x)\}, \{\psi^d_{j,2n+1}(x)\}$ are the respectively corresponding orthonormal bases of $V_{j,L}$ and $W_{j,L}$. Furthermore denoting $\{da_j(n)\}$ and $\{dd_j(n)\}$ as the coefficients of the decomposition of the details in frames $\{\psi^a_{j,p}(x)\}$ and $\{\psi^d_{j,p}(x)\}$, the following reconstruction relationship follows,

$$d_j(n) = (\sum_{-\infty}^{+\infty} da_j(2n)h(2n-p) + \sum_{-\infty}^{+\infty} dd_j(2n)g(2n-p)$$

or

$$d_j(n) = \sum_{-\infty}^{+\infty} da_j(2n+1)h(2n+1-p) + \sum_{-\infty}^{+\infty} dd_j(2n+1)g(2n+1-p).$$

Having established the above two lemmas and specializing it to a Haar function leads to

$$\phi(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{others}, \end{cases} \quad \psi(x) = \begin{cases} -1 & 0 < x < 0.5 \\ 1 & 0.5 < x < 1 \\ 0 & \text{otherwise}, \end{cases}$$

This leads to that the formula$\frac{1}{2}(u(t, i-1) - 2u(t, i) + u(t, i+1))$ is tantamount to iteratively subtracting details from detail to achieve a Heat diffusion and proves the proposition.

3.2. Haar Frame-Based Decomposition and Reconstruction

To further investigate the interplay and connection between PDE-based filtering and multiscale analysis, we proceed to specialize the foregoing development to a Haar wavelet frame and subsequently show an underlying direct connection to a Heat equation. For clarity of notation as well algebraic expediency, w adopt a matrix formalism which is also compatible with an image representation as a matrix. A nonorthogonal Haar representation of a signal can still yield a reconstruction. When using a Haar wavelet basis, the discrete form of the low pass filter is $h(0) = h(1) = \dfrac{1}{\sqrt{2}}$, and that of the high pass filter is $g(0) = -\dfrac{1}{\sqrt{2}}$ and $g(1) = \dfrac{1}{\sqrt{2}}$. First write the the following circulant matrices,

$$H = \mathrm{Cir}[h(0)\, h(1)] \,\mathrm{and} G = \mathrm{Cir}[g(0)\, g(1)]$$

Theorem 1. *Denote an initial image by a matrix A_0. Its redundant representation using a separable Haar function (i.e., obtaining the following spectral decomposition Low-Low,L0w-High, High-Low, High-High) can be written as*

$$\begin{aligned} A_1 = HA_0H'; \quad D_1 = HA_0G'; \\ D_2 = GA_0H'; \quad D_3 = GA_0G', \end{aligned}$$

where " ' " denotes transposition. The reconstruction matrices can similarly be written as

$$\begin{aligned} R^h_1 = h(0)I; \quad R^g_1 = g(0)I; \\ R^h_2 = h(1)I_{1,N}; \quad R^g_2 = g(1)I_{1,N} \end{aligned}$$

where

$$I_{1,N} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

in light of the fact that the redundant representation is given or may be computed, the exact reconstruction methods have to be carefully rewritten. Towards that end we have the following:

Proposition 2. *Any of the following four methods will exactly reconstruct the original image. Denoting,*

$$\begin{aligned} RA^{ij}_0 &= R^h_i A_1 R^{h\prime}_j ; RD^{ij}_1 &= R^h_i D_1 R^{g\prime}_j ; \quad (6) \\ RD^{ij}_2 &= R^g_i D_2 R^{h\prime}_j ; RD^{ij}_3 &= R^g_i D_3 R^{g\prime}_j , \quad i, j = 1, 2. \end{aligned}$$

we have,
method1: $A_0 = RA^{11}_0 + RD^{11}_1 + RD^{11}_2 + RD^{11}_3$
method2: $A_0 = RA^{21}_0 + RD^{21}_1 + RD^{21}_2 + RD^{21}_3$
method3: $A_0 = RA^{12}_0 + RD^{12}_1 + RD^{12}_2 + RD^{12}_3$
method4: $A_0 = RA^{22}_0 + RD^{22}_1 + RD^{22}_2 + RD^{22}_3$.

4. FRAME-BASED SMOOTHING

The above decomposition and reconstruction would be the same for Daubechies wavelet with a slight modification of the matrices H and G which are detailed in [9]. The smoothing in a frame representation may be carried out upon writing,

$$A_0 = RA^1_0 + R^h_1 D_1 R^{g\prime}_1 + R^g_2 D_2 R^{h\prime}_1 + R^g_1 D_3 R^{g\prime}_1 \quad (7)$$

and denoting the details of detail $D_i, i = 1, 2, 3, 4$ by $W^j_i, j = 1, 2, 3, 4$.. By using method 1-4 to reconstruct D_1, D_2, D_3, D_4, we use the knowledge that noise primarily dominates high frequency information. We then proceed to decrease its contribution by using the detail of detail terms from A_0 to result in the following recursion,

$$\begin{aligned} U_n &= U_{n-1} - (R^h_1 R^g_2 W^3_1 R^{g\prime}_2 R^{g\prime}_1 \\ &+ R^g_1 R^g_2 W^3_2 R^{g\prime}_2 R^{h\prime}_1 + R^g_1 R^g_2 W^3_3 R^{g\prime}_2 R^{g\prime}_1). \quad (8) \end{aligned}$$

Filtering via this recursion is equivalent to that of a Heat PDE, as show in prop.1. it behavior as a low pass filter which will ultimately completely smooth an image. Generalizing the above equation using higher order wavelets than Haar basis will achieve a more graceful lowpass filtering and a better preservation of features. Fig. 2 amplitudes of both low pass filters obtained by using Haar and Daubechies 4 demonstrate that additional high frequency content can pass.

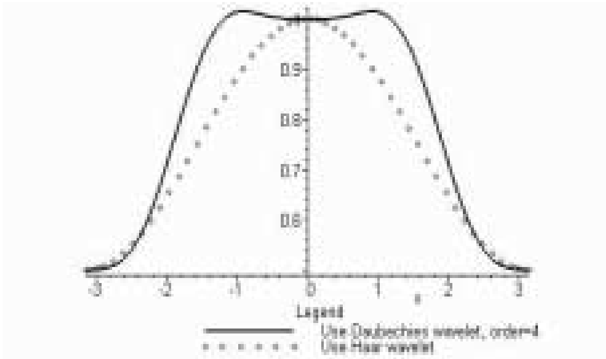

Figure 2: Compare Frequency content of low pass filter using Haar and Daubechies 4 basis.

4.1. Nonlinear Reconstruction

Inspired by the algorithms of the first section such as that we proposed or that of Perona-Malik, and to better preserve features well captured by some correlated structure reflected now in the coefficients, we can just as well proceed to construct a reconstruction filter based on a set of transformed frame coefficientsfrom Daubechies 4 as indicated below,
$D_1 = D_1 * exp(-D_1^2/2K)$;
$D_2 = D_2 * exp(-D_2^2/2K)$; $D_3 = D_3 * exp(-D_3^2/2K)$.
and iterate the reconstruction as noted above.

5. EXPERIMENTAL RESULTS AND CONCLUSION

To illustrate the performance of our proposed nonlinear filter, we show a Lenna picture with three different denoising techniques. Our originally proposed technique[10], Perona-Malik's, and the newly proposed technique. The advantages of removing noise while preserving features like texture are readily apparent in Figures 3, 4.

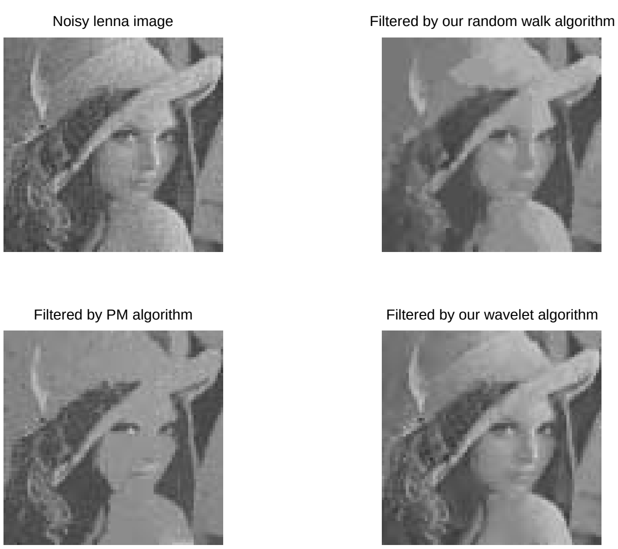

Figure 3: A noisy Lenna image and filtered result with three algorithms.

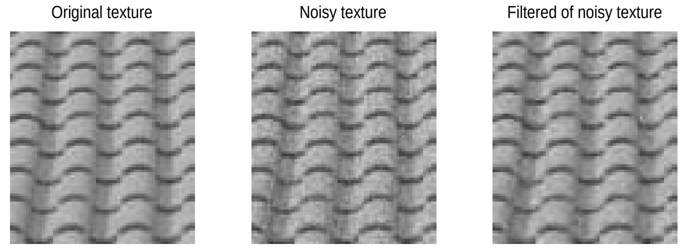

Figure 4: A texture image, noisy texture image and filtered result with Daubechies 4

6. REFERENCES

[1] A. P. Witkin, "Scale space filtering," in *Proc. Int. Joint Conf. on Artificial Intellingence*, (Karlsruhe, Germany), pp. 1019–1023, 1983.

[2] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-11, pp. 674–693, Jul. 1989.

[3] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage." preprint Dept. Stat., Stanford Univ., Jun. 1992.

[4] H. Krim and J.-C. Pesquet, *On the Statistics of Best Bases Criteria*, vol. Wavelets in Statistics of *Lecture Notes in Statistics*. Springer-Verlag, July 1995.

[5] N. Saito, *Local feature extraction and its applications using a library of bases*. PhD thesis, Yale University, Dec. 1994.

[6] P. Perona and J. Malik, "A network for multiscale image segmentation," in *IEEE Int. Symp. on Circuits and Systems*, (Helsinki), pp. 2565–2568, June 1988.

[7] B. ter Haar Romeney, *Geometry Driven Diffusion*. Dordecht: Kluwer Academic Publishers, 1994. Edited Book.

[8] I. Pollak, A. Willsky, and H. Krim, "Image segmentation and edge enhancement with stabilized inverse diffusion equations," *IEEE Trans. on Image Processing*, Feb, 2000.

[9] Y. Bao and H. Krim, "Nonlinear diffusion and wavelet denoising," *Journal paper. In Preparation*, 2001.

[10] H. Krim and Y. Bao, "Nonlinear diffusion: A probabilistic view," in *ICIP, Kobe, Japan, 1999.*

[11] S. Mallat, *A Wavelet Tour of Signal Processing*. Boston: Academic Press, 1997.