L.D. Griffin and M. Lillholm (Eds.): Scale-Space 2003, LNCS 2695, pp. 101-116, 2003.

Correspondences between Wavelet Shrinkage and Nonlinear Diffusion*

Pavel Mrázek[1], Joachim Weickert[1], and Gabriele Steidl[2]

[1] Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science, Building 27
Saarland University, 66123 Saarbrücken, Germany
{mrazek,weickert}@mia.uni-saarland.de
http://www.mia.uni-saarland.de

[2] Faculty of Mathematics and Computer Science, D7, 27
University of Mannheim, 68131 Mannheim, Germany
steidl@math.uni-mannheim.de
http://www.kiwi.math.uni-mannheim.de

Abstract. We study the connections between discrete one-dimensional schemes for nonlinear diffusion and shift-invariant Haar wavelet shrinkage. We show that one step of (stabilised) explicit discretisation of nonlinear diffusion can be expressed in terms of wavelet shrinkage on a single spatial level. This equivalence allows a fruitful exchange of ideas between the two fields. In this paper we derive new wavelet shrinkage functions from existing diffusivity functions, and identify some previously used shrinkage functions as corresponding to well known diffusivities. We demonstrate experimentally that some of the diffusion-inspired shrinkage functions are among the best for translation-invariant multiscale wavelet shrinkage denoising.

1 Introduction

We consider a classical task of signal denoising: create an estimate u of an original signal z from its noisy measurement f, where

$$f = z + n,$$

and n denotes an additive noise function. Various methods have been proposed to remove the noise from z without sacrificing important structures such as edges, including rank-order filtering, mathematical morphology, stochastic methods, adaptive smoothing, wavelet techniques, partial differential equations (PDEs)

* This joint research was supported by the project *Relations between nonlinear filters in digital image processing* within the DFG–Schwerpunktprogramm 1114: *Mathematical methods for time series analysis and digital image processing.* This is gratefully acknowledged.

and variational methods. Although these method classes serve the same purpose, relatively few publications examine their similarities and differences, in order to transfer results from one of these classes to the others, or to design hybrid methods that combine the advantages of different classes. The present paper is a contribution in this direction, where we concentrate on two of these methods, namely nonlinear diffusion techniques and wavelet shrinkage.

Nonlinear diffusion creates a family of restored signals $u(t)$ by starting from the noisy signal f, and evolving it locally according to a process described by a nonlinear partial differential equation. This process is controlled by a diffusivity function g of the signal gradient. Typically, $g(s)$ is a nonnegative, nonincreasing function of the gradient magnitude, approaching zero as $s \to \infty$. This setting leads to the effect that smoothing of u proceeds faster in homogeneous regions (where the gradient is small, caused possibly by noise), and discontinuities (large gradient, hopefully corresponding to important features of the underlying signal) tend to be preserved. Depending on the choice of the diffusivity function g, a single nonlinear diffusion equation may cover a variety of nonlinear filters, including the original nonlinear diffusion of Perona and Malik [27] and its regularised variants [8,31], total variation (TV) diffusion [2], or balanced forward-backward (BFB) diffusion [21]. When applied to discrete data $\mathbf{f} = (f_i)_{i=0}^{N-1}$, the nonlinear diffusion filter creates a series of smoothed signals $\mathbf{u}^k := \mathbf{u}(k\tau)$ iteratively, starting from the noisy signal, $\mathbf{u}^0 = \mathbf{f}$.

Wavelet transforms express the signal in terms of wavelet coefficients, describing the signal variation at different scales. If the wavelet basis is chosen properly, a signal will be generally described by only a few significant wavelet coefficients, while moderate white Gaussian noise pollutes all the wavelet coefficients by a small amount. Signal denoising by wavelet shrinkage [13,14] starts from this assumption, and creates a smoothed version of the processed signal by the following three-step procedure:

1. *Analysis:* transform the noisy data f to the wavelet coefficients d_i^j, representing the signal at various scales j and positions i.
2. *Shrinkage:* apply a shrinkage function S_θ to the wavelet coefficients d_i^j, thus reducing the relative importance of small coefficients.
3. *Synthesis:* reconstruct a denoised version u of f from the shrunken wavelet coefficients.

The shrinkage parameter θ is chosen with respect to the amount of noise in the input signal. In general, the denoised solution u is obtained from f using a single step of this multiscale procedure, i.e. the method is applied noniteratively. The specific choice of the wavelets and the shrinkage functions allows a large variability of wavelet shrinkage methods.

In the present paper, we show equivalence between a single iteration of a 1-D explicit scheme for nonlinear diffusion on one side, and translation-invariant wavelet shrinkage with a single level of Haar wavelet decomposition on the other. This equivalence is obtained by constructing an appropriate shrinkage function S_θ to an existing diffusivity g, and vice versa.

Having asserted the equivalence between wavelet shrinkage and nonlinear diffusion for this special situation, it remains to be seen whether this connection brings any advantages in more general settings. We demonstrate numerically that the shrinkage functions derived from diffusivities are able to provide some of the best results when used for classical (i.e. multi-level, one step) translation-invariant wavelet shrinkage.

This paper is organised as follows. Section 2 presents nonlinear diffusion and develops its explicit discretisation in 1-D. Section 3 provides a brief introduction into translation-invariant Haar wavelet shrinkage. The connections between the two procedures are exploited in Section 4 to establish the conditions on diffusivity and shrinkage functions under which the two methods (restricted to one-step / one-scale) are equivalent. Some newly created shrinkage function are then tested experimentally, and compared to previously used ones. The paper is concluded with a summary in Section 6.

Related work. Analysing the relations between regularisation methods and *continuous* wavelet shrinkage of functions, Chambolle *et al.* [5] showed that one may interpret wavelet shrinkage of functions as regularisation processes in suitable Besov spaces. In the case of Haar wavelets, Cohen *et al.* [9] showed that this approximates total variation regularisation. Later on, Chambolle and Lucier [6] considered iterated translation-invariant wavelet shrinkage and interpreted it as a nonlinear scale-space that differs from other scale-spaces by the fact that it is not given in terms of PDEs.

Regarding the relations between wavelet shrinkage denoising of *discrete* signals and nonlinear diffusion, not much research has been done so far. A recent paper by Coifman and Sowa [12] proposes TV diminishing flows that act along the direction of Haar wavelets. Recent work in which the authors are involved [29,30] investigates conditions under which equivalence between wavelet shrinkage of discrete signals, space-discrete TV diffusion or regularisation, and SIDEs (stabilised inverse diffusion equations) holds true.

Some recently proposed hybrid methods are based on combining wavelet shrinkage and TV regularisation methods [1,28]. Durand and Froment [15] proposed to address the problem of pseudo-Gibbs artifacts in wavelet denoising by replacing the thresholded wavelet coefficients by coefficients that minimise the total variation. Their method is also close in spirit to approaches by Chan and Zhou [7] who postprocessed images obtained from wavelet shrinkage by a TV-like regularisation technique. Coifman and Sowa [11] used functional minimisation with wavelet constraints for postprocessing signals that have been degraded by wavelet thresholding or quantisation. Candes and Guo [4] also presented related work, in which they combined ridgelets and curvelets with TV minimisation strategies. Recently, Malgouyres [23,24] proposed a hybrid method that uses both wavelet packets and TV approaches. His experiments showed that it may restore textured regions without introducing visible ringing artifacts.

This discussion shows that the previous papers typically focus on TV-based denoising techniques on the PDE side. Moreover, most of them present a continuous analysis rather than a discrete one. Our paper differs from previous work in this field by the fact that we do not restrict ourselves to a single diffusivity or shrinkage function, but introduce and analyse a general connection between a discrete diffusion scheme and Haar wavelet shrinkage. To this end, we investigate a large number of diffusivities and shrinkage functions.

2 Nonlinear Diffusion

2.1 Basic Concept

The basic idea behind nonlinear diffusion filtering [27] is to obtain a family $u(x,t)$ of filtered versions of the signal $f(x)$ as the solution of a suitable diffusion process

$$u_t = (g(|u_x|)\, u_x)_x \tag{1}$$

with f as initial condition:

$$u(x,0) = f(x).$$

Here subscripts denote partial derivatives, and the diffusion time t is a simplification parameter: larger values correspond to stronger filtering.

The diffusivity $g(|u_x|)$ is a nonnegative function that controls the amount of diffusion. Usually, it is decreasing in $|u_x|$. This ensures that strong edges are less blurred by the diffusion filter than noise and low-contrast details. Depending on the choice of the diffusivity function, equation (1) covers a variety of filters. Here are some of the previously employed diffusivity functions:

A. Linear diffusivity [19]: $g(|x|) = 1,$

B. Charbonnier diffusivity [8]: $g(|x|) = \dfrac{1}{\sqrt{1+\frac{|x|^2}{\lambda^2}}},$

C. Perona–Malik diffusivity [27]: $g(|x|) = \dfrac{1}{1+\frac{|x|^2}{\lambda^2}},$

D. Weickert diffusivity [31]: $g(|x|) = \begin{cases} 1 & |x| = 0, \\ 1-\exp\left(\frac{-3.31488}{(|x|/\lambda)^8}\right) & |x| > 0, \end{cases}$

E. TV diffusivity [2]: $g(|x|) = \dfrac{1}{|x|},$

F. BFB diffusivity: [21] $g(|x|) = \dfrac{1}{|x|^2}.$

Note that the diffusivities A–D are bounded from above by 1, while the diffusivities E and F are unbounded. In order to avoid theoretical and numerical difficulties, it is common to replace the latter ones by regularisations that make them bounded: e.g. one may use $g(|x|) = 1/\sqrt{\epsilon^2 + |x|^2}$ instead of the TV diffusivity.

Well-posedness results are available for the diffusivities A, B and E, since they lead to forward parabolic processes. For the diffusivities C, D and F, which may lead to backward parabolic equations, well-posedness questions are open in the continuous setting [22,20], while a space-discretisation seems to lead to well-posed processes [32].

2.2 Explicit Discretisation Scheme

When applied to discrete signals, the partial differential equation (1) has to be discretised. In this paper, we focus on explicit finite difference schemes. Substituting the spatial partial derivatives in (1) by finite differences (with the assumption of unit distance between neighboring pixels), and employing explicit discretisation in time, an explicit 1-D scheme for nonlinear diffusion can be written in the form

$$\frac{u_i^{k+1}-u_i^k}{\tau} \;=\; g(|u_{i+1}^k-u_i^k|)\,(u_{i+1}^k-u_i^k)-g(|u_i^k-u_{i-1}^k|)\,(u_i^k-u_{i-1}^k),$$

where τ is the time step size and the upper index k denotes the approximate solution at time $k\tau$. Separating the unknown u_i^{k+1} on one side, we obtain

$$u_i^{k+1}=u_i^k-\tau\,g(|u_i^k-u_{i+1}^k|)\,(u_i^k-u_{i+1}^k)+\tau\,g(|u_{i-1}^k-u_i^k|)\,(u_{i-1}^k-u_i^k). \quad (2)$$

The initial condition reads $u_i^0=f_i$ for all i.

3 Wavelet Shrinkage

3.1 Basic Concept

The discrete wavelet transform represents a one-dimensional signal f in terms of shifted versions of a dilated lowpass scaling function φ, and shifted and dilated versions of a bandpass wavelet function ψ. In case of orthonormal wavelets, this gives

$$f=\sum_{i\in\mathbb{Z}}\langle f,\varphi_i^n\rangle\,\varphi_i^n+\sum_{j=-\infty}^{n}\sum_{i\in\mathbb{Z}}\langle f,\psi_i^j\rangle\,\psi_i^j, \quad (3)$$

where $\psi_i^j(s):=2^{-j/2}\psi(2^{-j}s-i)$ and where $\langle\cdot,\cdot\rangle$ denotes the inner product in $L_2(\mathbb{R})$. If the measurement f is corrupted by moderate white Gaussian noise, then this noise is contained to a small amount in all wavelet coefficients $\langle f,\psi_i^j\rangle$, while the original signal is in general determined by a few significant wavelet coefficients [25]. Therefore, wavelet shrinkage attempts to eliminate noise from the wavelet coefficients by the following three-step procedure:

1. *Analysis*: transform the noisy data f to the wavelet coefficients $d_i^j=\langle f,\psi_i^j\rangle$ and scaling function coefficients $c_i^n=\langle f,\varphi_i^n\rangle$ according to (3).
2. *Shrinkage*: apply a shrinkage function S_θ with a threshold parameter θ to the wavelet coefficients, i.e., $S_\theta(d_i^j)=S_\theta(\langle f,\psi_i^j\rangle)$.

3. *Synthesis*: reconstruct the denoised version u of f from the shrunken wavelet coefficients:

$$u := \sum_{i\in\mathbb{Z}} \langle f, \varphi_i^n\rangle\, \varphi_i^n + \sum_{j=-\infty}^{n} \sum_{i\in\mathbb{Z}} S_\theta(\langle f, \psi_i^j\rangle)\, \psi_i^j .$$

In this paper we restrict our attention to Haar wavelets, well suited for piecewise constant signals with discontinuities. The Haar wavelet and scaling functions are given respectively by

$$\psi(x) = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}, \tag{4}$$

$$\phi(x) = \mathbf{1}_{[0,1)} \tag{5}$$

where $\mathbf{1}_{[a,b)}$ denotes the characteristic function, equal to 1 on $[a, b)$ and zero everywhere else. Using the so-called "two-scale relation" of the wavelet and its scaling function, the coefficients c_i^j and d_i^j at higher level j can be computed from the coefficients c_i^{j-1} at lower level $j-1$ and conversely:

$$c_i^j = \frac{c_{2i}^{j-1} + c_{2i+1}^{j-1}}{\sqrt{2}}, \qquad d_i^j = \frac{c_{2i}^{j-1} - c_{2i+1}^{j-1}}{\sqrt{2}}, \tag{6}$$

and

$$c_{2i}^{j-1} = \frac{c_i^j + d_i^j}{\sqrt{2}}, \qquad c_{2i+1}^{j-1} = \frac{c_i^j - d_i^j}{\sqrt{2}}. \tag{7}$$

This results in a fast algorithm for the analysis step and synthesis step. Various shrinkage functions leading to qualitatively different denoised functions u were considered in literature, e.g.,

A. Linear shrinkage: $S(x) = \lambda x \qquad (\lambda \in [0,1])$,

B. Soft shrinkage [13]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x - \theta\ \mathrm{sgn}(x) & |x| > \theta, \end{cases}$

C. Garrote shrinkage [16]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x - \frac{\theta^2}{x} & |x| > \theta, \end{cases}$

D. Firm shrinkage [17]: $S_{\theta_1,\theta_2}(x) = \begin{cases} 0 & |x| \le \theta_1, \\ \mathrm{sgn}(x)\frac{\theta_2(|x|-\theta_1)}{\theta_2-\theta_1} & \theta_1 < |x| \le \theta_2, \\ x & \theta_2 < |x|, \end{cases}$

E. Hard shrinkage [25]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x & |x| > \theta. \end{cases}$

3.2 Discrete Translation-Invariant Scheme

In practice one deals with discrete signals $\mathbf{f} = (f_i)_{i=0}^{N-1}$, where, for simplicity, N is a power of 2. Then Haar wavelet shrinkage starts by setting $c_i^0 = f_i$ and proceeds by analysis (6), shrinkage, and synthesis (7). Let us just consider a *single* wavelet decomposition level, i.e., we set $n = 1$. Then, using the convention that $c_i = c_i^1$ and $d_i = d_i^1$, we can drop the superscripts $j = 0$ and $j = 1$. By (6) and (7), Haar wavelet shrinkage on one level produces the signal $\mathbf{u}^+ = (u_i^+)_{i=0}^{N-1}$ with coefficients

$$u_{2i}^+ = \frac{c_i + S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right), \tag{8}$$

$$u_{2i+1}^+ = \frac{c_i - S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right). \tag{9}$$

Note that the single Haar wavelet shrinkage step (8)–(9) decouples the input signal into successive pixel pairs: the pixel at position $2i-1$ has no direct connection to its neighbour at position $2i$, and the procedure is not invariant to translation of the input signal. To overcome this problem, Coifman and Donoho [10] introduced the so-called *cycle spinning*: the input signal is shifted, denoised using wavelet shrinkage, shifted back, and the results of all such shifts are averaged. This procedure is equivalent to thresholding of nondecimated wavelet coefficients which can be implemented efficiently using the *algorithme à trous* [18]. For our single decomposition level, we need only one additional shift to acquire translation invariance. The shifted Haar wavelet shrinkage yields the signal $\mathbf{u}^- = (u_i^-)_{i=0}^{N-1}$ with coefficients

$$\begin{aligned} u_{2i-1}^- &= \frac{f_{2i-1} + f_{2i}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1} - f_{2i}}{\sqrt{2}}\right), \\ u_{2i}^- &= \frac{f_{2i-1} + f_{2i}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1} - f_{2i}}{\sqrt{2}}\right). \end{aligned}$$

Averaging the shifted results, one cycle of shift-invariant Haar wavelet shrinkage can be summarised into

$$\begin{aligned} u_i &= \frac{u_i^- + u_i^+}{2} \\ &= \frac{f_{i-1} + 2f_i + f_{i+1}}{4} + \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_i - f_{i+1}}{\sqrt{2}}\right) - \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_{i-1} - f_i}{\sqrt{2}}\right). \end{aligned} \tag{10}$$

4 Correspondence of Diffusivities and Shrinkage Functions

4.1 Basic Considerations

In order to derive the relation between the explicit diffusion scheme and translation-invariant Haar wavelet shrinkage, we rewrite the first iteration step in (2) using the initial condition $u_i^0 = f_i$ and the simplified notation $u_i^1 = u_i$ as

$$
\begin{aligned}
u_i = {} & \frac{f_{i-1} + 2f_i + f_{i+1}}{4} + \frac{f_i - f_{i+1}}{4} - \frac{f_{i-1} - f_i}{4} \\
& - \tau\, g(|f_i - f_{i+1}|)\,(f_i - f_{i+1}) + \tau\, g(|f_{i-1} - f_i|)\,(f_{i-1} - f_i) \\
= {} & \frac{f_{i-1} + 2f_i + f_{i+1}}{4} \\
& + (f_i - f_{i+1}) \left(\frac{1}{4} - \tau\, g(|f_i - f_{i+1}|) \right) \\
& - (f_{i-1} - f_i) \left(\frac{1}{4} - \tau\, g(|f_{i-1} - f_i|) \right). \qquad (11)
\end{aligned}
$$

This coincides with (10) if and only if

$$
\frac{1}{2\sqrt{2}}\, S_\theta\!\left(\frac{x}{\sqrt{2}} \right) = x \left(\frac{1}{4} - \tau\, g(|x|) \right). \qquad (12)
$$

Equation (12) relates the shrinkage function S_θ of wavelet denoising to the diffusivity g of nonlinear diffusion. Provided that relation (12) holds true, a single step of wavelet shrinkage is equivalent to a single step of explicitly discretised nonlinear diffusion. The following two formulas are derived from (12) and can be used to obtain a shrinkage function S_θ from a diffusivity g, or vice versa.

$$
S_\theta(x) = x \left(1 - 4\tau g(|\sqrt{2}x|) \right), \qquad (13)
$$

$$
g(|x|) = \frac{1}{4\tau} - \frac{\sqrt{2}}{4\tau x}\, S_\theta\!\left(\frac{x}{\sqrt{2}} \right). \qquad (14)
$$

4.2 From Diffusivities to Shrinkage Functions

Let us now investigate equation (13) in detail. The examples from Section 3.1 show that typical shrinkage functions from the literature satisfy

$$
S(x) \geq 0 \quad \text{for} \quad x > 0, \qquad (15)
$$

$$
S(x) \leq 0 \quad \text{for} \quad x < 0. \qquad (16)
$$

One can show that these conditions are responsible for ensuring certain stability properties (so-called sign stability) of the shrinkage process. We can now specify the time step size τ in (13) such that these two conditions are always satisfied for bounded diffusivities. In Section 2.1 we have seen that the diffusivities A–D are bounded by 1. In order to ensure that the corresponding shrinkage functions satisfy (15)–(16), the time step size has to fulfil $\tau \leq 0.25$.

We observe that the linear diffusivity corresponds to the linear shrinkage function

$$
S(x) = (1 - 4\tau)x.
$$

Nonlinear shrinkage functions such as soft, garrote, firm and hard shrinkage satisfy $S'(0) = 0$, since the goal was to set small wavelet coefficients to zero. In order to derive shrinkage functions that correspond to the bounded nonlinear diffusivities B–D and satisfy $S'(0) = 0$ as well, let us now fix $\tau := 0.25$. Then we obtain the following novel shrinkage functions:

- The Charbonnier diffusivity corresponds to the shrinkage function

$$S_\lambda(x) = x\left(1 - \sqrt{\frac{\lambda^2}{\lambda^2 + 2x^2}}\right).$$

- The Perona–Malik diffusivity leads to

$$S_\lambda(x) = \frac{2x^3}{2x^2 + \lambda^2}.$$

- The Weickert diffusivity gives

$$S_\lambda(x) = \begin{cases} 0 & x = 0, \\ x\,\exp\left(-\frac{0.20718\,\lambda^8}{x^8}\right) & x \neq 0. \end{cases}$$

Figure 1 illustrates these bounded diffusivities and their shrinkage functions.

4.3 From Shrinkage Functions to Diffusivities

Having derived shrinkage functions from nonlinear diffusivities, let us now derive diffusivities from frequently used shrinkage functions. To this end, all we have to do is to plug in the specific shrinkage function into (14).
In the case of soft shrinkage, this gives the diffusivity

$$g(|x|) = \begin{cases} \frac{1}{4\tau} & |x| \leq \theta\sqrt{2}, \\ \frac{\sqrt{2}\theta}{4\tau|x|} & |x| > \theta\sqrt{2}. \end{cases}$$

If we select the time step size τ such that $\theta = 2\sqrt{2}\tau$, we obtain a stabilised TV diffusivity:

$$g(|x|) = \begin{cases} \frac{1}{4\tau} & |x| \leq 4\tau, \\ \frac{1}{|x|} & |x| > 4\tau. \end{cases}$$

In the same way one can show that garrote shrinkage leads to a stabilised BFB diffusivity for $\theta = \sqrt{2\tau}$:

$$g(|x|) = \begin{cases} \frac{1}{4\tau} & |x| \leq 2\sqrt{\tau}, \\ \frac{1}{|x|^2} & |x| > 2\sqrt{\tau}. \end{cases}$$

Firm shrinkage yields a diffusivity that degenerates to 0 for sufficiently large gradients:

$$g(|x|) = \begin{cases} \frac{1}{4\tau} & |x| \leq \sqrt{2}\theta_1, \\ \frac{\theta_1}{4\tau(\theta_2-\theta_1)}\left(\frac{\sqrt{2}\theta_2}{|x|} - 1\right) & \sqrt{2}\theta_1 < |x| \leq \sqrt{2}\theta_2, \\ 0 & |x| > \sqrt{2}\theta_2. \end{cases}$$

Such diffusivities have been considered in [3], where they have been motivated using priors from robust statistics.
Another diffusivity that degenerates to 0 can be derived from hard shrinkage:

$$g(|x|) = \begin{cases} \frac{1}{4\tau} & |x| \leq \sqrt{2}\theta, \\ 0 & |x| > \sqrt{2}\theta. \end{cases}$$

All diffusivities in this subsection are depicted in Figure 2.

5 Denoising Experiment

To test the applicability of the newly derived shrinkage functions from Subsection 4.2, we perform experiments with signal-denoising using the shift-invariant multiscale Haar wavelet transform from Section 3. The input signal *blocks*, one of the standard signals in wavelet denoising, mimics a scan line through a 2-D image depicting an object with several edges [14]. The signal is shown in Fig. 3. The same figure then shows examples of the results of multiscale Haar wavelet denoising when combined with several shrinkage functions introduced in previous sections.

SNR_{in} / Shrinkage method	1	2	4	8	16	32
Linear	3.6	4.2	5.5	8.7	16.1	32.0
Soft (TV)	10.0	10.8	12.6	16.2	24.0	39.9
Perona-Malik	9.9	10.8	12.7	16.8	25.8	44.6
Weickert	12.7	13.7	15.9	20.3	29.4	46.3
Garrote (BFB)	11.8	12.8	14.9	19.3	28.7	46.2
Firm	12.6	13.6	15.8	20.2	29.3	46.3
Hard	12.7	13.8	15.9	20.4	29.3	46.3

Table 1. Numerical results (measured by mean signal-to-noise ratio in the filtered signal) of wavelet denoising for the *blocks* data of length 1024. Each column represents a given level of noise in the input image; each row contains the results for one shrinkage function.

Table 1 and Fig. 4 present additional experimental results obtained with the *blocks* data. Here we performed a series of experiments with several levels of additive zero-mean Gaussian noise in the input signal. The noise varies between SNR=1 and SNR=32, where the signal-to-noise ratio (SNR) is defined by $\text{SNR} = 20 \log_{10} \frac{|z-\bar{z}|_2}{|n|_2}$, with z standing for the ideal signal with mean $\bar{z}$, and n representing noise. The noise is generated five times for each SNR level. Then we used multiscale wavelet denoising with various shrinkage functions, and searched for the optimal solution that can be obtained with this method. By optimal we mean the solution maximising the signal-to-noise ratio in the filtered signal.

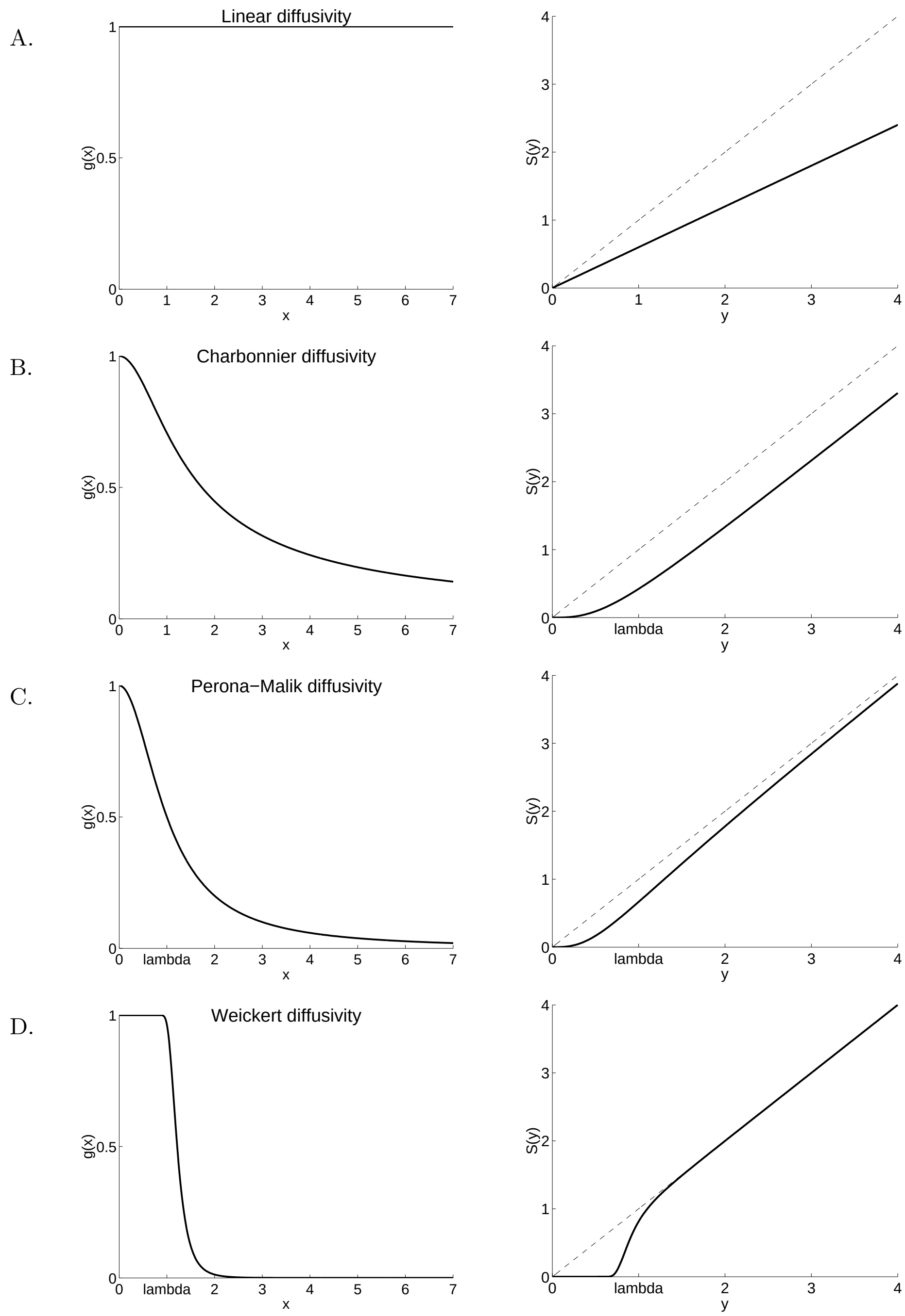

Fig. 1. Diffusivity functions (left), corresponding shrinkage functions (right). A. Linear diffusion. B. Charbonnier diffusivity. C. Perona-Malik diffusivity. D. Weickert diffusivity. The functions are plotted for $\tau = 0.1$ (linear diffusion), and $\tau = 0.25$, $\lambda = 1$ (all others).

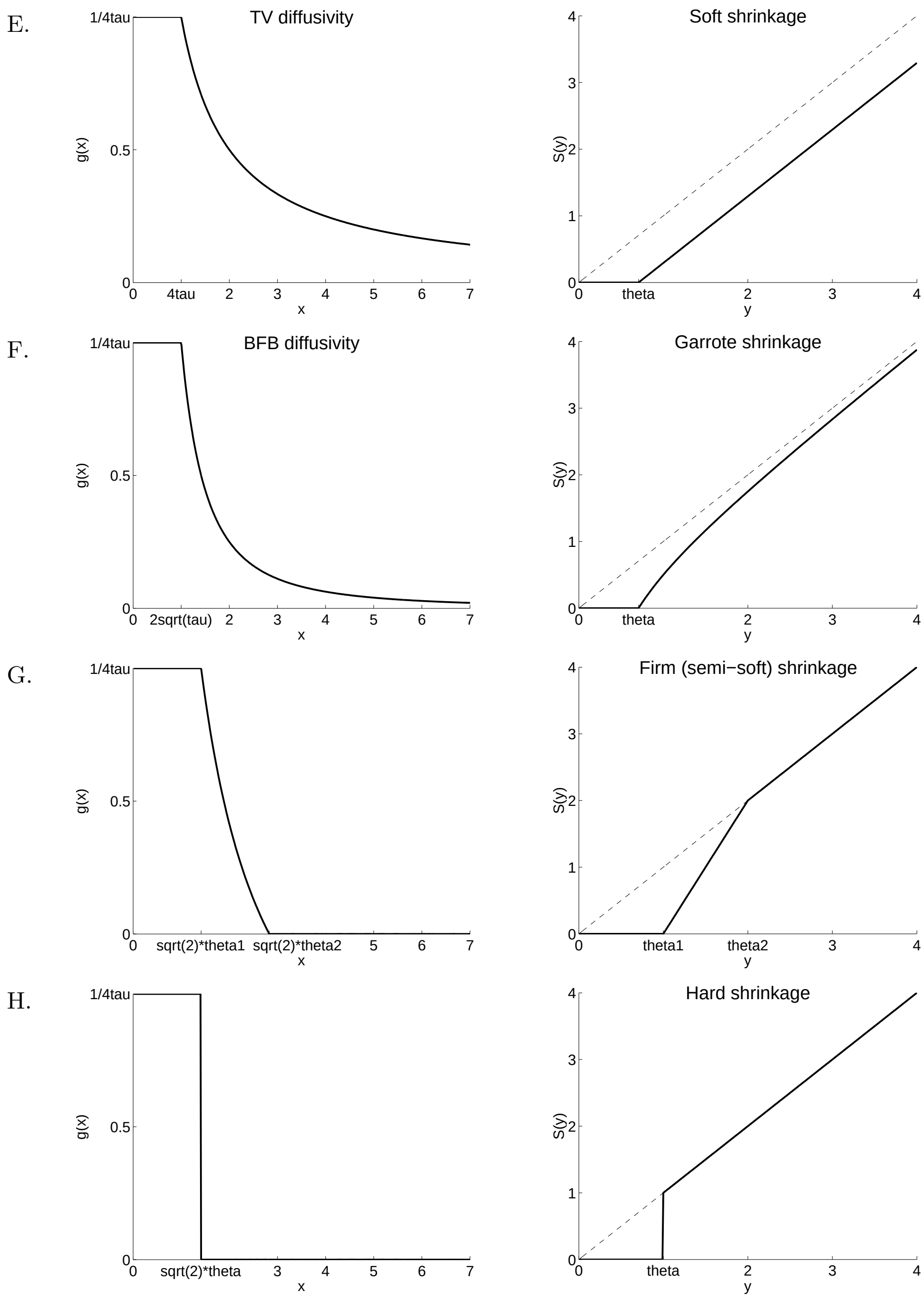

Fig. 2. Diffusivity functions (left), corresponding shrinkage functions (right). E. TV flow and soft shrinkage. F. Balanced forward-backward (BFB) diffusivity and garrote shrinkage. G. Firm shrinkage. H. Hard shrinkage.
The functions are plotted for $\tau = 0.25$ (which corresponds to $\theta = \tau 2\sqrt{2}$ for soft shrinkage, and to $\theta = \sqrt{2\tau}$ for garrote; the other use $\theta = \theta_1 = 1$, $\theta_2 = 2$).

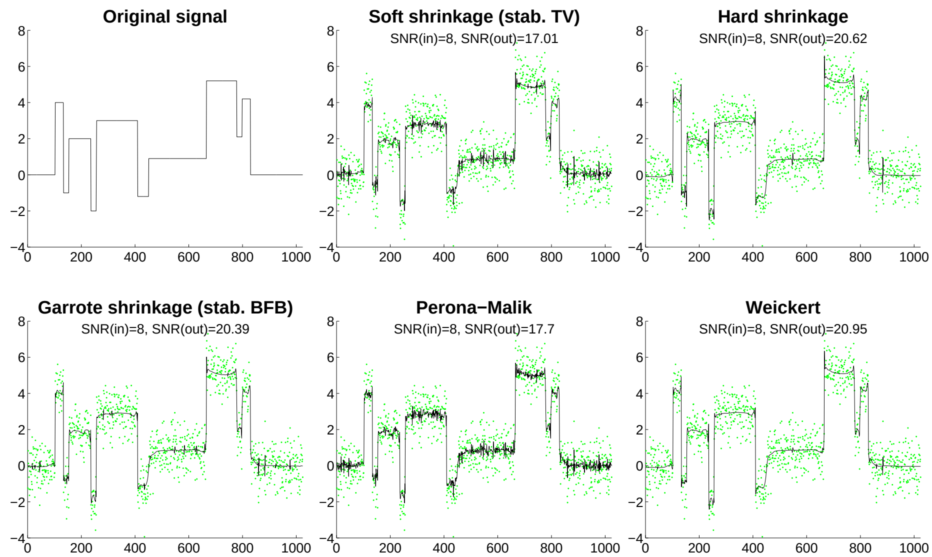

Fig. 3. Example of multiscale translation-invariant Haar wavelet denoising. Normal noise of SNR=8 was added to the ideal signal, and different shrinkage functions have been applied. The noisy signal is represented by dots, reconstructed signal by solid line.

Table 1 summarises the average optimal SNR after filtering obtained with different shrinkage functions; Fig. 4 presents the same information graphically, together with the standard deviation of the results. We observe that for all noise levels, the best signal-to-noise ratio is obtained by those shrinkage functions which put small wavelet coefficients to zero and keep larger coefficients almost unaffected. The functions with these properties include hard shrinkage, firm shrinkage and – to some extent – the garrote shrinkage on the wavelet side. Of the diffusion origin, the experimentally best shrinkage functions correspond to Weickert diffusivity, stabilised BFB diffusivity (which is equivalent to garrote shrinkage), and Perona-Malik diffusivity. Interestingly, these are diffusivities with nonmonotone flux functions that allow even contrast enhancement.

The second group of shrinkage functions decreases even large wavelet coefficients by a constant (or almost constant) value; the functions with this behaviour include soft shrinkage, TV flow corresponding to it, and Charbonnier diffusivity. It seems that this strategy is less successful numerically. These diffusivities lead to monotonically increasing flux functions and well-posed diffusion filters.

As a group of its own, the denoising performance of linear diffusion (or its shrinkage function) is far worse than that of the nonlinear methods.

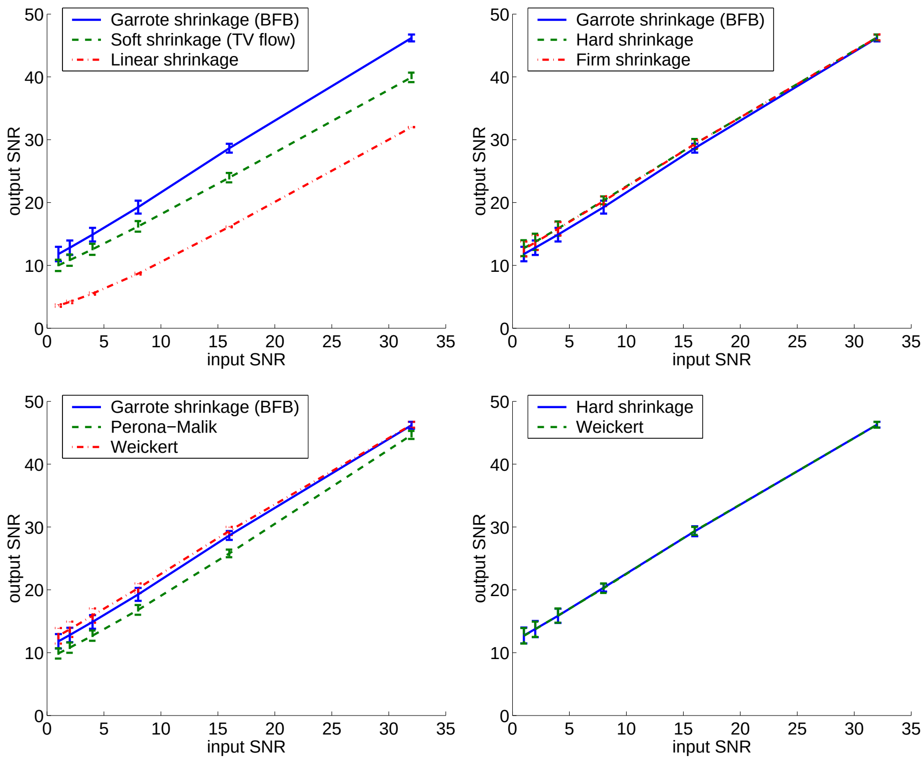

Fig. 4. Comparing the optimal denoising performance of shift-invariant multi-scale wavelet shrinkage with various shrinkage functions. SNR of the filtered signal is plotted against SNR of the input; the higher the graph, the better the result. The input signal was *blocks*, length 1024.
Top left: garrote shrinkage (BFB diffusivity), soft shrinkage (TV flow) and linear diffusion. Top right: garrote (BFB), hard and firm shrinkages. Bottom left: garrote (BFB), Perona-Malik and Weickert functions. Bottom right: best from either world, hard shrinkage and Weickert diffusivity give comparable results.

6 Conclusions

We have analysed correspondences between explicit one-dimensional schemes for nonlinear diffusion and discrete translation-invariant Haar wavelet shrinkage. We have shown that if we restrict the methods to one discrete step and a single spatial level, the two approaches can be made equivalent, if suitable diffusivities or shrinkage functions are chosen.

This connection between nonlinear diffusion and wavelet shrinkage opens the gate for a fruitful exchange of ideas between the two worlds. In this paper, we derived new wavelet shrinkage functions from frequently used nonlinear diffusivities; vice versa, we showed that soft and garrote shrinkage may be regarded as stabilised TV or BFB diffusion, respectively. We experienced that the novel shrinkage functions corresponding to rapidly decreasing diffusivities are competitive with the best previously known shrinkage methods when applied to signal denoising with multiscale wavelet procedures.

The results in this paper can be extended in several directions. One can study iterated multi-scale wavelet shrinkage as a hybrid method combining the efficiency of multi-scale wavelet shrinkage with the quality of iterated diffusion filtering [26]. This hybrid method may be also explained as nonlinear diffusion applied to the Laplacian pyramid of the signal [29,30]. In our ongoing work, we are considering other wavelet bases and the two-dimensional case.

References

1. R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill–posed problems. *Inverse Problems*, 10:1217–1229, 1994.
2. F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential and Integral Equations*, 14(3):321–360, March 2001.
3. M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE Transactions on Image Processing*, 7(3):421–432, March 1998.
4. E. J. Candés and F. Guo. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Signal Processing*, 82(11):1519–1543, 2002.
5. A. Chambolle, R. A. DeVore, N. Lee, and B. L. Lucier. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7(3):319–335, March 1998.
6. A. Chambolle and B. L. Lucier. Interpreting translationally-invariant wavelet shrinkage as a new image smoothing scale space. *IEEE Transactions on Image Processing*, 10(7):993–1000, 2001.
7. T. F. Chan and H. M. Zhou. Total variation improved wavelet thresholding in image compression. In *Proc. Seventh International Conference on Image Processing*, Vancouver, Canada, September 2000.
8. P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proc. 1994 IEEE International Conference on Image Processing*, volume 2, pages 168–172, Austin, TX, November 1994. IEEE Computer Society Press.
9. A. Cohen, R. DeVore, P. Petrushev, and H. Xu. Nonlinear approximation and the space $BV(R^2)$. *Americal Journal of Mathematics*, 121:587–628, 1999.
10. R. R. Coifman and D. Donoho. Translation invariant denoising. In A. Antoine and G. Oppenheim, editors, *Wavelets in Statistics*, pages 125–150. Springer, New York, 1995.
11. R. R. Coifman and A. Sowa. Combining the calculus of variations and wavelets for image enhancement. *Applied and Computational Harmonic Analysis*, 9(1):1–18, July 2000.
12. R. R. Coifman and A. Sowa. New methods of controlled total variation reduction for digital functions. *SIAM Journal on Numerical Analysis*, 39(2):480–498, 2001.
13. D. L. Donoho. De-noising by soft thresholding. *IEEE Transactions on Information Theory*, 41:613–627, 1995.
14. D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrica*, 81(3):425–455, 1994.

15. S. Durand and J. Froment. Reconstruction of wavelet coefficients using total-variation minimization. Technical Report 2001–18, Centre de Mathématiques et de Leurs Applications, ENS de Cachan, France, 2001.
16. H.-Y. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7(4):469–488, 1998.
17. H.-Y. Gao and A. G. Bruce. WaveShrink with firm shrinkage. *Statistica Sinica*, 7:855–874, 1997.
18. M. Holschneider, R. Kronland-Martinet, J. Morlet, and Ph. Tchamitchian. A real-time algorithm for signal analysis with the help of the wavelet transform. In J.M. Combes, A. Grossman, and Ph. Tchamitchian, editors, *Wavelets: Time-Frequency Methods and Phase Space*, pages 286–297. Springer-Verlag, 1987.
19. T. Iijima. Basic theory on normalization of pattern (in case of typical one-dimensional pattern). *Bulletin of the Electrotechnical Laboratory*, 26:368–388, 1962. In Japanese.
20. B. Kawohl and N. Kutev. Maximum and comparison principle for one-dimensional anisotropic diffusion. *Mathematische Annalen*, 311:107–123, 1998.
21. S. L. Keeling and R. Stollberger. Nonlinear anisotropic diffusion filters for wide range edge sharpening. *Inverse Problems*, 18:175–190, January 2002.
22. M. Kijima. *Markov Processes and Stochastic Modeling.* Chapman and Hall, New York, 1997.
23. F. Malgouyres. Combining total variation and wavelet packet approaches for image deblurring. In *Proc. First IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 57–64, Vancouver, Canada, July 2001. IEEE Computer Society Press.
24. F. Malgouyres. Mathematical analysis of a model which combines total variation and wavelet for image restoration. *Inverse Problems*, 2(1):1–10, 2002.
25. S. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, San Diego, second edition, 1999.
26. Pavel Mrázek, Joachim Weickert, Gabriele Steidl, and Martin Welk. On iterations and scales of nonlinear filters. In O. Drbohlav, editor, *Computer Vision Winter Workshop 2003*, pages 61–66. Czech Pattern Recognition Society, 2003.
27. P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.
28. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
29. G. Steidl and J. Weickert. Relations between soft wavelet shrinkage and total variation denoising. In L. Van Gool, editor, *Pattern Recognition*, volume 2449 of *Lecture Notes in Computer Science*, pages 198–205. Springer, Berlin, 2002.
30. G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. Technical report, Series SPP-1114, Department of Mathematics, University of Bremen, Germany, 2003.
31. J. Weickert. *Anisotropic Diffusion in Image Processing.* Teubner, Stuttgart, 1998.
32. J. Weickert and B. Benhamouda. A semidiscrete nonlinear scale-space theory and its relation to the Perona–Malik paradox. In F. Solina, W. G. Kropatsch, R. Klette, and R. Bajcsy, editors, *Advances in Computer Vision*, pages 1–10. Springer, Wien, 1997.