Inverse Problems and Imaging
Volume 1, No. 1, 2007, 47–62
Web site: http://www.aimSciences.org

INTEGRODIFFERENTIAL EQUATIONS FOR CONTINUOUS MULTISCALE WAVELET SHRINKAGE

Stephan Didas and Joachim Weickert

Department of Mathematics and Computer Science
Saarland University, Building E1 1
66041 Saarbrücken, Germany

Abstract. The relations between wavelet shrinkage and nonlinear diffusion for discontinuity-preserving signal denoising are fairly well-understood for single-scale wavelet shrinkage, but not for the practically relevant multiscale case. In this paper we show that 1-D multiscale continuous wavelet shrinkage can be linked to novel integrodifferential equations. They differ from nonlinear diffusion filtering and corresponding regularisation methods by the fact that they involve smoothed derivative operators and perform a weighted averaging over all scales. Moreover, by expressing the convolution-based smoothed derivative operators by power series of differential operators, we show that multiscale wavelet shrinkage can also be regarded as averaging over pseudodifferential equations.

1. Introduction

Wavelet shrinkage and nonlinear diffusion filtering constitute two important classes of discontiuity-preserving denoising methods for signals and images. The research on both filtering techniques started in the early 1990s, and meanwhile the proposed methods are relatively well understood and often used in practice.

During the last years there has been a growing interest in analysing the relations between wavelet-based methods and methods based on partial differential equations (PDEs) such as diffusion filters and their corresponding variational approaches [1]. This includes both works in the continous [2, 6, 7, 10, 11, 15, 28, 40] and in the discrete setting [13, 29, 30, 41, 46]. Let us now focus on some examples that are particularly relevant in the context of the present paper.

In the continuous setting, a variational formulation for wavelet shrinkage has been presented by Chambolle et al. [6, 7]. This makes it possible to understand wavelet shrinkage as an image smoothing scale-space. Since the scale-space concept originated from diffusion equations [23, 48], this already points out some structural similarities. Nevertheless, a direct comparison between the scale-space properties remained still open. Bredies et al. [3] show the equivalence of variational wavelet shrinkage to abstract pseudodifferential evolution equations. One open question in practice is the connection of these pseudodifferential equations to well-known image processing methods such as nonlinear diffusion of Perona-Malik type.

2000 *Mathematics Subject Classification:* 68U10, 47G20, 65T60.
Key words and phrases: image processing, integrodifferential equations, wavelet shrinkage, nonlinear diffusion filtering.

[1] Sometimes wavelet shrinkage and PDE-based denoising methods are also used in combination [4, 8, 12, 20, 21, 26].

Equivalence results between methods of both classes have also been shown for discrete problems and under certain conditions [41, 30]. For example, discrete soft Haar wavelet shrinkage on the finest scale is equivalent to total variation flow and total variation diffusion [41]. Under more general assumptions, for other diffusivities or with other kinds of wavelets, one still can show close relationships [30, 45]. These ideas are based on the fact that discrete wavelets on the finest scale approximate derivatives. The derivative order is determined by the number of vanishing moments of the discrete filters. An open question in this context is what happens if we take coarser scales than only the finest one into consideration. In this context only experimental results are available [31]: They indicate that iterations within nonlinear diffusion implementations play a similar role as performing shift-invariant wavelet shrinkage at multiple scales. So far, however, no results have been derived that may be helpful in understanding the differences between both techniques.

The goal of the present paper is to address the before mentioned open questions. To this end we consider wavelet shrinkage in the practically relevant *multiscale* setting. For the sake of simplicity we focus on 1-D signals and we analyse the continuous shrinkage framework only. The key observation that we exploit in our paper is the fact that, for wavelets with a finite number of vanishing moments, the wavelet transform can be understood as applying a smoothed derivative operator [27, p. 167]. Thus, wavelet shrinkage on a single scale is closely related to a diffusion type equation where all appearing spatial derivatives are regularised with a convolution kernel. Going from a single scale to multiple scales introduces a further integration, yielding a novel integrodifferential equation derived from wavelet shrinkage. Moreover, we express convolution with a smoothing kernel by a power series of differential operators. This allows us to regard multiscale wavelet shrinkage as an averaging of pseudodifferential equations over a continuum of scales. These results make the analytical reasons for the differences between continuous multiscale wavelet shrinkage on one side and nonlinear diffusion and its corresponding variational regularisation on the other side explicit: They are caused by the presence of additional integration scales, smoothing operators and differential operators of higher order.

Our paper is organised as follows: Section 2 introduces some useful notations and summarises the classical continuous wavelet shrinkage approach, while Section 3 gives a brief introduction of diffusion filtering techniques. The idea of understanding wavelets as smoothed derivative operators, which is crucial for the remainder of the paper, is explained in Section 4. With this knowledge, Section 5 describes how wavelet shrinkage can be interpreted as approximation to a novel integrodifferential evolution equation. In Section 6 we are going to present a corresponding energy functional that uses both smoothed derivative operators within the penaliser and integration over all scales. The link to pseudodifferential operators is discussed in Section 7. Section 8 concludes the paper with a summary and sketches some problems for ongoing and future research.

2. Wavelet Shrinkage

This section sketches the ideas behind wavelet shrinkage in a mathematical formulation that is suitable for our further considerations.

Wavelet shrinkage became popular by the work of Donoho and Johnstone [18]. The general idea is to transform the data to a representation that allows to reduce noise in a straightforward way, namely by diminishing the modulus of the wavelet coefficients. Especially the low computational complexity of the wavelet transform has made such approaches highly interesting for signal and image processing applications. As shown in [6, 7, 3], continuous wavelet shrinkage can be understood as minimisation of certain energy functionals. Of special interest has been soft wavelet shrinkage, which is related to gradient descent along Besov norms.

Before we formulate continuous wavelet shrinkage, let us give an introduction to the formal notions related to wavelets which will be used in this paper. As usual let $L^p(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid \int_{\mathbb{R}} |f(x)|^p \, dx < \infty\}$ for $1 \leq p < \infty$. In this and the following three sections, we consider signals as real functions $f, u \in L^1(\mathbb{R})$. We choose a real function $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, the *mother wavelet*, which has to satisfy the *admissibility condition* [14, p. 27]

$$c_\psi \ := \ 2\pi \int_0^\infty \frac{|\hat{\psi}(\xi)|^2}{\xi} \, d\xi \ < \ \infty \ . \tag{1}$$

Here, the Fourier transform is defined as

$$F\psi(\xi) \ := \ \hat{\psi}(\xi) \ = \ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) \exp(-ix\xi) \, dx \ . \tag{2}$$

To simplify the notation let ψ_σ and $\tilde{\psi}$ be scaled and mirrored versions of ψ, i.e. $\psi_\sigma(x) := \frac{1}{\sqrt{c_\psi \sigma}} \psi\left(\frac{x}{\sigma}\right)$ and $\tilde{\psi}(x) := \psi(-x)$. By $f * g$ we denote the convolution of two functions $f, g \in L^1(\mathbb{R})$:

$$(f * g)(x) \ = \ \int_{-\infty}^{\infty} f(x - \tau) g(\tau) \, d\tau \quad \text{for all } x \in \mathbb{R} \ . \tag{3}$$

If a function $f(y, z)$ depends on more than one variable we replace the variable in which the convolution is performed with a dot, for example

$$(f(\cdot, z) * g)(x) \ = \ \int_{\mathbb{R}} f(x - \tau, z) g(\tau) \, d\tau \quad \text{for all } x, z \in \mathbb{R} \ . \tag{4}$$

This notations help us to write down wavelet shrinkage in an easy way. Wavelet shrinkage transforms the data in a suitable representation and performs simple nonlinear operations. The back-transform finally yields the denoising result. Let us now formulate these three steps in detail:

1. **Analysis:** First the given function f is transformed into the wavelet domain. With the notations introduced above we can write the wavelet transform as

$$W_\psi f(x, \sigma) \ := \ \frac{1}{\sqrt{c_\psi}} \int_{-\infty}^{\infty} f(\tau) \frac{1}{\sqrt{\sigma}} \psi\left(\frac{\tau - x}{\sigma}\right) \, d\tau \ = \ \left(\tilde{\psi}_\sigma * f\right)(x) \ . \tag{5}$$

   Thus this step can be seen as convolution with scaled and mirrored versions of the mother wavelet.
2. **Shrinkage:** A – typically nonlinear – shrinkage function $S : \mathbb{R} \longrightarrow \mathbb{R}$ is applied to the wavelet transform $W_\psi f$. Usually it is assumed that this shrinkage function diminishes the absolute value of the wavelet coefficients without changing their sign. Reasonable assumptions on S are thus

$$x > 0 \Longrightarrow S(x) \geq 0 \ , \quad S(-x) = -S(x) \ , \quad \text{and } |S(x)| \leq |x| \tag{6}$$

   for all $x \in \mathbb{R}$. Usually S depends on a parameter λ which determines the amount of shrinkage. This parameter is omitted here to simplify the notations. Examples of shrinkage functions can be found in Tab. 1.

3. **Synthesis:** As final step, the shrunken wavelet transform $S \circ W_\psi f$ has to be transformed back into the spatial domain to yield the resulting function u of the shrinkage. It should be mentioned that the wavelet transform W is an isometric map from $L^2(\mathbb{R})$ to $L^2(\mathbb{R} \times \mathbb{R}_0^+)$, equipped with the inner product

$$\langle f, g \rangle := \int_{-\infty}^{\infty} \int_0^{\infty} f(x,\sigma)\, \bar{g}(x,\sigma)\, \frac{d\sigma}{\sigma^2}\, dx \tag{7}$$

where $\bar{g}$ is the complex conjugate of g. Thus, on the subspace given by the image of W_ψ, the adjoint operator W_ψ^* is the inverse of W_ψ. We show several ways to formulate the back-transform:

$$\begin{aligned} u &= W_\psi^*(S \circ W_\psi f) \\ &= \int_0^{\infty} \Big(\psi_\sigma * S(W_\psi f(\cdot,\sigma))\Big)\, \frac{d\sigma}{\sigma^2} \\ &= \frac{1}{\sqrt{c_\psi}} \int_{-\infty}^{\infty} \int_0^{\infty} S(W_\psi f(\tau,\sigma))\, \frac{1}{\sqrt{\sigma}}\, \psi\left(\frac{\cdot - \tau}{\sigma}\right)\, \frac{d\sigma}{\sigma^2}\, d\tau \ . \end{aligned} \tag{8}$$

The convergence of these integrals should be understood in the weak sense, see [14, p.25]. In the following sections we will mostly refer to the following formulation:

$$u = \int_0^{\infty} \Big(\psi_\sigma * S\big(\tilde{\psi}_\sigma * f\big)\Big)\, \frac{d\sigma}{\sigma^2} \ . \tag{9}$$

Note that besides the convolution with ψ_σ, the back-transform also introduces an integration over all scales σ.

The formulation for wavelet shrinkage given in this section will be the starting point for our considerations in the following.

3. Nonlinear Diffusion Filtering and Regularisation

The goal of this section is to sketch nonlinear diffusion processes of arbitrary order and their corresponding regularisation methods, and to consider modificiations that turn the differential equations into integrodifferential equations. All these processes are required when interpreting multiscale wavelet shrinkage in terms of an intergrodifferential equation.

3.1. Nonlinear Diffusion Filtering of Arbitrary Order.

Nonlinear diffusion filtering in image processing goes back to the seminal paper of Perona and Malik [35] in 1990; for an overview we refer to [44].

One of the major drawbacks of diffusion filters is the so-called staircasing effect: Especially edge-enhancing processes tend to turn smooth grey value transitions into stair-like artifacts in the image. One way to circumvent this problem is to introduce higher derivative orders in the filtering process [49, 25, 16]. This generalises piecewise constant results to piecewise polynomial ones which may be better adapted to certain applications.

In this paper we consider higher order nonlinear diffusion filters with the following family of PDEs:

$$\partial_t u = (-1)^{n+1}\, \partial_x^n \left(g\left(|\partial_x^n u|^2\right) \partial_x^n u \right) \ . \tag{10}$$

For $n \in \mathbb{N} \setminus \{0\}$ we call this a *diffusion filter of order* $2n$ because this is the maximal derivative order appearing in the equation. We notice that the case $n = 1$ yields the classical model by Perona and Malik. For practical purposes usually the cases $n = 1$

and $n = 2$ are most interesting. Orders $n > 2$ often tend to adapt too much on high noise outliers. Furthermore, in a 2-D grey value image, the difference between a quadratic and a cubic polynomial region is optically hardly distinguishable.

3.2. CORRESPONDING REGULARISATION METHODS. It is well-known that nonlinear diffusion filtering can be closely related to regularisation methods [38]. Such regularisation strategies with first order derivatives in image processing have been considered for example in [34, 9, 39, 32]. Higher derivative orders are also used for regularisation purposes, see [37]. Such variational approaches obtain denoised or simplified versions of the data by minimising energy functionals of the form

$$E(u) \;=\; \int_{-\infty}^{\infty} \left((u-f)^2 + \alpha\, p\left(|u^{(n)}|^2 \right) \right)\, dx \quad . \tag{11}$$

The function p in this case is an increasing function which penalises large fluctuations in term of pronounced derivatives of order n. The choice of this penalising function heavily influences the characteristics of the solutions.

The connection of (11) to the nonlinear diffusion equation (10) becomes evident when we write down the Euler-Lagrange equation of (11):

$$\frac{u-f}{\alpha} \;=\; (-1)^{n+1}\, \frac{d^n}{dx^n} \left(p'\left(|u^{(n)}|^2 \right) u^{(n)} \right) \quad . \tag{12}$$

This equation can be regarded as a fully implicit time discretisation of (10) with initial condition $u(x,0) = f(x)$, diffusivity $g = p'$ and time step size α.

3.3. MODIFICATIONS LEADING TO INTEGRODIFFERENTIAL EQUATIONS. Besides the pure regularisation and diffusion models discussed so far, we will now point out with several examples that it is common practice in image processing to use smoothing kernels in combination with derivatives: A prominent example for such a filter is the regularised nonlinear diffusion by Catté et al. [5]. It is a second-order Perona-Malik model where the data steering the diffusivity is smoothed with a Gaussian kernel before taking the derivative:

$$\partial_t u \;=\; \partial_x \Big(g\left(|\partial_x G_\sigma * u|^2 \right)\, \partial_x u \Big) \tag{13}$$

In this example, the presmoothing introduces well-posedness in the problem and renders the filter more robust against noise.

Other authors proposed to replace all derivative operators consequently with Gaussian-smoothed derivatives [33]. An example which would fit into this strategy is regularisation with a presmoothed derivative [38]:

$$E(u) \;=\; \int_{-\infty}^{\infty} \left((u-f)^2 + \alpha\, p\left(\left| \frac{d}{dx}\, G_\sigma * u \right|^2 \right) \right)\, dx \quad . \tag{14}$$

The corresponding diffusion process to this energy functional is given by

$$\partial_t u \;=\; \partial_x\, G_\sigma * \left(g\left(|\partial_x\, G_\sigma * u|^2 \right) \partial_x\, G_\sigma * u \right). \tag{15}$$

However, these approaches have hardly found practical applications so far. One problem is that smoothing all derivatives with a fixed scale σ makes the method incapable of removing noise on a smaller scale. This cannot happen with the method by Catté et al. since the outer derivatives are not smoothed.

As a last example in this context we want to mention edge-enhancing anisotropic diffusion filtering in the case of at least two-dimensional data [44]:

$$\partial_t u \;=\; \operatorname{div} \left(D(\nabla\, G_\sigma * u)\, \nabla u \right). \tag{16}$$

This filter family is of special interest as example because presmoothing is necessary for the anisotropy of the filter: Due to the construction of the diffusion tensor D, the preferred directions for the anisotropy are the one of $\nabla G_\sigma * u$ and its orthogonal direction. Without presmoothing one would simply go in direction of ∇u which directly leads to an isotropic model.

These examples show that presmoothing of derivatives within diffusion filters is a common concept that is applied in different contexts and for several reasons. Let us now investigate its interpretation in terms of the wavelet transform.

4. Wavelet Transforms as Smoothed Derivative Operators

In Section 2 we have introduced the wavelet shrinkage technique with an arbitrary mother wavelet ψ. Now we restrict our choice to a certain class of wavelets in order to relate the corresponding wavelet transform to smoothed derivative operators.

First we assume that the mother wavelet ψ has fast decay, i. e. for any exponent $m \in \mathbb{N}$ there exists a constant c_m such that

$$|\psi(x)| \;\leq\; \frac{c_m}{1+|x|^m} \quad \text{for all } x \in \mathbb{R} \;. \tag{17}$$

For the rest of the paper, we focus on wavelets with a finite number $n \in \mathbb{N} \setminus \{0\}$ of vanishing moments:

$$\int_{-\infty}^{\infty} x^k \psi(x)\; dx = 0 \quad \text{for } 0 \leq k < n \;. \tag{18}$$

It is well-known [27, p.167] that these assumptions are equivalent to the existence of a function θ with fast decay such that

$$\psi(x) \;=\; (-1)^n \, \frac{d^n}{dx^n}\, \theta(x) \;. \tag{19}$$

Moreover, ψ has no more than n vanishing moments if and only if θ has nonzero mean value, i. e. $\int_{-\infty}^{\infty} \theta(x)\; dx \neq 0$. In our further considerations, this function θ will play the role of a smoothing kernel. Keeping this in mind it makes sense to consider the maximal number of vanishing moments for a certain wavelet which gives a natural choice of the corresponding derivative order.

Fig. 1 gives two examples of wavelets that are often used in image processing together with their corresponding derivative orders and smoothing kernels. It can be seen that the Haar wavelet is the first derivative of a hat-shaped function. The second classical example is the Mexican hat wavelet which is defined as the second derivative of a Gaussian kernel. Examples of wavelets with a higher number of vanishing moments include the Daubechies wavelets. All the wavelet classes mentioned here are also of fast decay, the Mexican hat wavelet because of its exponential decreasing velocity, and the others because of their compact support.

So far, we have seen the relation (19) between the mother wavelet and the smoothing kernel. It is not difficult to verify the following equations for scaled and mirrored versions of the wavelet:

$$\tilde{\psi}_\sigma * f \;=\; \sigma^n \, \partial_x^n \left(\tilde{\theta}_\sigma * f\right) \;=\; \sigma^n \, (\tilde{\theta}_\sigma)^{(n)} * f \tag{20}$$

and

$$\psi_\sigma * f \;=\; (-\sigma)^n \, \partial_x^n \,(\theta_\sigma * f) \;=\; (-\sigma)^n \,(\theta_\sigma)^{(n)} * f \;. \tag{21}$$

As an elementary property of the convolution, one can also put the derivative in front of the function f, if the regularity of f allows for this. If this is not the case

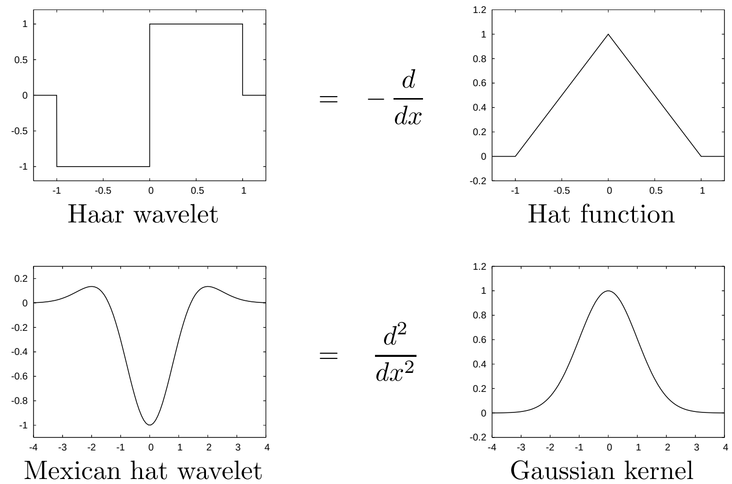

FIGURE 1. Examples of wavelets as derivatives of smoothing kernels.

one can motivate the smoothing kernel θ as regularisation of the n-th derivative of f.

Equation (20) shows that the wavelet transform is equivalent to taking a smoothed derivative with an additional weight factor σ^n:

$$W_\psi f \;=\; \left(\tilde{\psi}_\sigma * f\right) \;=\; \sigma^n\, \partial_x^n \left(\tilde{\theta}_\sigma * f\right)(x) \;. \tag{22}$$

For the back-transform, an additional integration over all scales σ is introduced:

$$W_\psi^* f \;=\; \int_0^\infty (\psi_\sigma * f)\; \frac{d\sigma}{\sigma^2} \;=\; \int_0^\infty (-\sigma)^n\, \partial_x^n\, (\theta_\sigma * f)\; \frac{d\sigma}{\sigma^2} \;. \tag{23}$$

These two equations will form the basis for relating wavelet techniques to derivative-based methods.

5. WAVELET SHRINKAGE AND EVOLUTION EQUATIONS

In this section, we are going to relate wavelet shrinkage as given in Section 2 to novel integrodifferential evolution equations involving the data on a continuous spectrum of scales.

The prototype of an evolution equation in the sense of this paper can be written down informally as

$$\partial_t u \;=\; L^*\Big(g\left(|Lu|^2\right) Lu\Big) \;. \tag{24}$$

In principle we will study this in the case that L is the wavelet transform. Additionally we are going to argue that the time-discrete shrinkage steps can be seen just as an approximation to a time-continuous evolution.

We start with writing down one step of wavelet shrinkage as given in Section 2:

$$u \;=\; \int_0^\infty \Big(\psi_\sigma * S\big(\tilde{\psi}_\sigma * f\big)\Big)\; \frac{d\sigma}{\sigma^2} \;. \tag{25}$$

Now we choose a function g such that

$$S(x) \;=\; x - g(|x|^2)\, x \;. \tag{26}$$

Shrinkage	$S(x)$	$g(\lvert x\rvert^2)$	Diffusivity
Linear	λx	$1-\lambda$	Linear [23]
Hard [27]	$\begin{cases} 0, & \lvert x\rvert \le \lambda \\ x, & \lvert x\rvert > \lambda \end{cases}$	$\begin{cases} 1, & \lvert x\rvert \le \lambda \\ 0, & \lvert x\rvert > \lambda \end{cases}$	
Soft [17]	$\begin{cases} 0, & \lvert x\rvert \le \lambda \\ x-\lambda\,\mathrm{sgn}(x), & \lvert x\rvert > \lambda \end{cases}$	$\begin{cases} 1, & \lvert x\rvert \le \lambda \\ \frac{\lambda}{\lvert x\rvert}, & \lvert x\rvert > \lambda \end{cases}$	$\approx$ TV [1, 36]
Garrote [22]	$\begin{cases} 0, & \lvert x\rvert \le \lambda \\ x-\frac{\lambda^2}{x}, & \lvert x\rvert > \lambda \end{cases}$	$\begin{cases} 1, & \lvert x\rvert \le \lambda \\ \frac{\lambda^2}{x^2}, & \lvert x\rvert > \lambda \end{cases}$	$\approx$ BFB [24]
	$\frac{x^3}{\lambda^2+x^2}$	$\left(1+\frac{x^2}{\lambda^2}\right)^{-1}$	Perona-Malik [35]
	$\left(1-\sqrt{\frac{\lambda^2}{\lambda^2+x^2}}\right)x$	$\left(1+\frac{x^2}{\lambda^2}\right)^{-\frac{1}{2}}$	Charbonnier [9]

TABLE 1. Shrinkage functions and corresponding diffusivities. Adapted from [30].

During the further steps we are going to see immediately why this choice makes sense (cf. also [30] where a relation of this type has been derived in a discrete setting). The function g is going to play the role of a diffusivity function in the following considerations. Under the assumptions (6) it follows that $0 \le g(|x|^2) \le 1$. This range is in accordance with stability requirements that do not allow negative diffusivities, and bounding the diffusivity from above by 1 is a frequently used normalisation in image analysis.

Tab. 1 shows some corresponding pairs of shrinkage functions S and diffusivities g. We obeserve that several practically used shrinkage functions lead to well-known diffusivities. An interesting case is for example soft wavelet shrinkage [17] that corresponds via (26) to a total variation (TV) diffusivity [1] which is regularised for small parameters where it would become unbounded. A similar regularisation for small parameters appears in the balanced-forward-backward (BFB) diffusivity [24] corresponding to Garrote shrinkage. With the Perona-Malik diffusivity [35] and the diffusivity related to the regularisation approach of Charbonnier et al. [9] we also give two examples where we start with classical diffusivities and calculate corresponding shrinkage functions. Plugging (26) into the wavelet shrinkage formula (25) yields

$$u = \int_0^\infty \left(\psi_\sigma * \tilde{\psi}_\sigma * f\right) \frac{d\sigma}{\sigma^2} - \int_0^\infty \left(\psi_\sigma * \left(g\left(|\tilde{\psi}_\sigma * f|^2\right)(\tilde{\psi}_\sigma * f)\right)\right) \frac{d\sigma}{\sigma^2}\ . \tag{27}$$

The first integral is simply the wavelet reconstruction formula which gives back the initial data f. Therefore this expression is equivalent to

$$u - f = -\int_0^\infty \left(\psi_\sigma * \left(g\left(|\tilde{\psi}_\sigma * f|^2\right)(\tilde{\psi}_\sigma * f)\right)\right) \frac{d\sigma}{\sigma^2}\ . \tag{28}$$

The next step now is to involve the ideas from Section 4 to regard wavelets as smoothed derivative operators. We obtain the following equivalent formulation of wavelet shrinkage as integrodifferential equation:

$$\boxed{u - f \; = \; (-1)^{n+1} \int_0^\infty \sigma^{2n} \, \partial_x^n \, \theta_\sigma * \left(g \left(|\sigma^n \, \partial_x^n \, \tilde{\theta}_\sigma * f|^2 \right) (\partial_x^n \, \tilde{\theta}_\sigma * f) \right) \; \frac{d\sigma}{\sigma^2} \; .} \tag{29}$$

Let us now relate this process to nonlinear diffusion filtering. Similar to the approach in [38] we now introduce an artificial time variable t for the function u such that the initial data is given at time $t = 0$: $u(\cdot, 0) = f$. It is possible to write the left-hand side as $\frac{u(x)-f(x)}{\tau} = \frac{u(x,\tau)-u(x,0)}{\tau}$ with $\tau = 1$. Keeping this in mind, (28) can be understood as a single step time-explicit approximation to the evolution equation

$$\boxed{\partial_t u = (-1)^{n+1} \int_0^\infty \sigma^{2n} \, \partial_x^n \, \theta_\sigma * \left(g \left(|\sigma^n \, \partial_x^n \, \tilde{\theta}_\sigma * u|^2 \right) (\partial_x^n \, \tilde{\theta}_\sigma * u) \right) \; \frac{d\sigma}{\sigma^2} \; .} \tag{30}$$

If we compare (30) to the higher order nonlinear diffusion equation

$$\partial_t u = (-1)^{n+1} \, \partial_x^n \left(g \left(|\partial_x^n u|^2 \right) \partial_x^n u \right) \tag{31}$$

we notice two differences:

1. All appearing derivative operators in the equation are pre-smoothed by convolution with scaled and mirrored versions of a kernel θ.
2. The right-hand side is not only considered at one single scale, but there is an integration over all scales with additional weight factors.

We are going to see in the next section that these two steps can also be used to turn classical variational methods into wavelet-based ones.

It should be noted that the choice $\tau = 1$ in the above derivation is motivated by the fact that we want to represent exactly one step of wavelet shrinkage with this stopping time. By using the relation

$$S(x) = (1 - \tau \tilde{g}(|x|^2))x \tag{32}$$

instead of (26), one could specify an arbitrary stopping time $\tau > 0$ which corresponds to one shrinkage step. This would lead to other diffusivity functions $\tilde{g} = \frac{g}{\tau}$. We have chosen $\tau = 1$ here to avoid the division by τ in all diffusivities.

We have seen how wavelet shrinkage can be related to integrodifferential equations. In the next section we consider regularisation approaches and investigate how these can be related to variational formulations for wavelet shrinkage.

6. Variational Formulation and Correspondences

Our prototype of a variational method for denoising or simplifying a signal f can be written as follows:

$$E(u) \; = \; \int_{-\infty}^{\infty} (u - f)^2 \; dx \, + \, \alpha \int_\Gamma p \left(|Lu(y)|^2 \right) \; dy \; . \tag{33}$$

First a linear operator L is used to extract some features of interest from the image u which are penalised with a typically nonlinear function p. The variable name y should indicate that the function Lu can live in another domain than u which is named Γ here. Examples for such operators L can be derivative operators, the wavelet transform, or also the Fourier transform.

Here we consider that case that L is the wavelet transform as introduced in Section 2, so we have $\Gamma = \mathbb{R} \times \mathbb{R}_0^+$ and integrate with the corresponding weight. Thus we formulate an energy functional for wavelet shrinkage as

$$\begin{aligned} E(u) &= \int_{-\infty}^{\infty} (u-f)^2 \, dx + \alpha \int_{-\infty}^{\infty} \int_0^{\infty} p\left(|W_\psi u(x,\sigma)|^2\right) \frac{d\sigma}{\sigma^2} \, dx \\ &= \int_{-\infty}^{\infty} (u-f)^2 \, dx + \alpha \int_{-\infty}^{\infty} \int_0^{\infty} p\left(|\tilde{\psi}_\sigma * u|^2\right) \frac{d\sigma}{\sigma^2} \, dx \ . \end{aligned} \tag{34}$$

This is closely connected to ideas and variational formulations presented in [6, 7, 3].

With (20) we can rewrite this energy functional as

$$E(u) = \int_{-\infty}^{\infty} (u-f)^2 \, dx + \alpha \int_{-\infty}^{\infty} \int_0^{\infty} p\left(|\sigma^n \, \partial_x^n (\tilde{\theta}_\sigma * u)|^2\right) \frac{d\sigma}{\sigma^2} \, dx \ . \tag{35}$$

Again we see that the same two steps as in the last section lead to classical regularisation approaches (11). First we leave out the integration over all scales and obtain a variational functional

$$E(u) = \int_{-\infty}^{\infty} (u-f)^2 \, dx + \alpha \int_{-\infty}^{\infty} p\left(|\partial_x^n \, (\tilde{\theta}_\sigma * u)|^2\right) \, dx \ . \tag{36}$$

A special case ($n = 1$) of this functional is given by (14) which has been considered by Scherzer and Weickert [38]. In our case a richer choice of convolution kernels and higher derivative orders is allowed. The second step is to omit also the convolution. This leads us directly to the classical regularisation functionals (11).

By comparing these results with the last section, we see that the two major differences multiscale wavelet shrinkage and regularisation are the same as between multiscale wavelet shrinkage and nonlinear diffusion: smoothed derivative operators instead of derivatives, and weighted integration over all scales instead of working at the finest scale.

7. Smoothing Kernels and Pseudodifferential Operators

So far, we have related wavelet shrinkage to integrodifferential equations which involve convolutions with scaled and mirrored versions of the mother wavelet. In this section, we express these convolutions as pseudodifferential operators. This allows us to eliminate all integrals from (30) except the integration over all scales. Instead of the convolutions, the equations then contain power series of differential operators.

After introducing some notations and technical details, we are going to describe the general procedure. Then we apply this to examples for convolution kernels, namely to a box function and a Gaussian kernel. As we have seen in Section 2, also the wavelet transform is a convolution and thus considered after the first two examples. Since the Haar and the Mexican hat wavelet can be written as derivatives of a box function and a Gaussian, we can conclude the section with these two popular examples of wavelets.

Let $\mathcal{S}$ be the Schwartz space of rapidly decreasing functions [42, 43]. We consider the convolution $\theta * f$ of a function $f \in \mathcal{S}(\mathbb{R})$ with a kernel $\theta \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. It is well-known that convolution in the spatial domain is equivalent to multiplication in the Fourier domain:

$$F(h * f) = \sqrt{2\pi} \, \hat{h} \cdot \hat{f} \ . \tag{37}$$

Besides the convolution, also derivative operators are multiplications in the Fourier domain, namely

$$F\left(\frac{d}{dx}f\right) \;=\; i\xi\hat{f} \;. \tag{38}$$

Let us assume that the Fourier transform of our convolution kernel θ is analytic, i. e. there is a power series representation

$$\hat{\theta}(\xi) \;=\; \sum_{k=0}^{\infty} a_k\, \xi^k \;. \tag{39}$$

For $f \in \mathcal{S}(\mathbb{R})$ we can understand the product $\hat{\theta}\cdot\hat{f}$ as a sum of derivatives of f up to arbitrary orders in the Fourier domain:

$$\frac{1}{\sqrt{2\pi}}\,\theta * f \;=\; F^{-1}(\hat{\theta}\cdot\hat{f}) \;=\; F^{-1}\left(\sum_{k=0}^{\infty} a_k\,\xi^k\,\hat{f}\right) \;. \tag{40}$$

Under the assumption of sufficient convergence conditions of the power series, we may interchange the sum and the Fourier back-transform which allows us to write

$$F^{-1}\left(\sum_{k=0}^{\infty} a_k\,\xi^k\,\hat{f}\right) \;=\; \sum_{k=0}^{\infty} a_k\,F^{-1}(\xi^k\hat{f}) \;=\; \sum_{k=0}^{\infty} a_k\left(\frac{1}{i}\,\frac{d}{dx}\right)^k f \;. \tag{41}$$

In this context the symbol $\hat{\theta}\left(\frac{1}{i}\partial_x\right)$ is used to denote this power series of differential operators (see [43], for example):

$$\hat{\theta}\left(\frac{1}{i}\frac{d}{dx}\right)f \;:=\; \sum_{k=0}^{\infty} a_k\left(\frac{1}{i}\frac{d}{dx}\right)^k f \;=\; \frac{1}{\sqrt{2\pi}}\,\theta * f \;. \tag{42}$$

We reformulate this as the central equation

$$\boxed{\theta * f \;=\; \sqrt{2\pi}\;\hat{\theta}\left(\frac{1}{i}\,\frac{d}{dx}\right)f} \tag{43}$$

relating convolution to power series of derivatives. It is also well-known that such a reasoning can be generalised from analytic functions to richer classes of functions $\hat{\theta}$, for example continuous or measurable functions [42, 47].

After describing the general idea, let us now apply this to two examples of convolution kernels:

Convolution with a box function. A basic operation in image processing is to take the arithmetic mean inside a symmetric neighbourhood of pixels. This can be understood as convolution with the characteristic function of an interval

$$\chi(x) \;:=\; \chi_{[-\frac{1}{2},\frac{1}{2}]}(x) \;=\; \begin{cases} 1, & x \in [-\frac{1}{2},\frac{1}{2}] \\ 0, & \text{else} \;. \end{cases} \tag{44}$$

This function has the Fourier transform

$$\hat{\chi}(\xi) = \frac{1}{\sqrt{2\pi}}\operatorname{sinc}\left(\frac{1}{2}\xi\right) = \frac{1}{\sqrt{2\pi}}\,\frac{\sin\left(\frac{1}{2}\xi\right)}{\frac{1}{2}\xi} = \frac{1}{\sqrt{2\pi}}\sum_{k=0}^{\infty}\frac{(-1)^k\xi^{2k}}{4^k(2k+1)!} \;. \tag{45}$$

Thus we write the convolution with the box function χ as pseudodifferential operator

$$\chi * f \;=\; \operatorname{sinc}\left(\frac{1}{2i}\frac{d}{dx}\right)f \;=\; \sum_{k=0}^{\infty}\frac{(-1)^k}{4^k(2k+1)!}\,\frac{d^{2k}}{dx^{2k}}\,f \;. \tag{46}$$

Convolution with a Gaussian kernel. As our second example, we take a look at the convolution with a Gaussian kernel, since this operation is fundamental in image processing. It is well-known [43] that the Fourier transform of a Gaussian kernel $\theta(x) = \exp\left(-\frac{x^2}{2}\right)$ is again a Gaussian function $\hat{\theta}(\xi) = \exp\left(-\frac{\xi^2}{2}\right)$. With the above reasoning we then write

$$\theta * f \;=\; \sqrt{2\pi}\,\exp\left(\frac{1}{2}\frac{d^2}{dx^2}\right) f \;=\; \sqrt{2\pi}\sum_{k=0}^{\infty}\frac{1}{2^k k!}\,\frac{d^{2k}}{dx^{2k}}\,f \;\;. \tag{47}$$

Similar formulations often appear in the context of linear scale-spaces and the heat equation (see [32, 19, 43]).

After these two examples, we now apply the idea to the convolution coming from the wavelet transform.

Wavelet transform. We have seen in Section 2 how the wavelet transform can be understood as convolution operator. Given a mother wavelet ψ, the Riemann-Lebesgue theorem assures that its Fourier transform is integrable, i. e. $\hat{\psi} \in L^1(\mathbb{R})$. This allows us to use (43) and write

$$\psi * f \;=\; \sqrt{2\pi}\;\hat{\psi}\left(\frac{1}{i}\,\frac{d}{dx}\right) f \;\;. \tag{48}$$

For a wavelet transform we now need convolutions with translated and scaled versions of the mother wavelet ψ which can be obtained in a general way. The translation is simply the evaluation of $\psi * f$ at another point. For the scaled and mirrored version a substitution shows that

$$(F\tilde{\psi}_\sigma)(\xi) \;=\; -\sqrt{\sigma}\,\hat{\psi}(-\sigma\xi) \;\;. \tag{49}$$

Together this means that we can write the wavelet transform as

$$W_\psi f(x,\sigma) \;=\; \tilde{\psi}_\sigma * f(x) \;=\; -\sqrt{2\pi\sigma}\left(\hat{\psi}\left(-\frac{\sigma}{i}\frac{d}{dx}\right) f\right)(x) \tag{50}$$

and thus express it as pseudodifferential operator.

As we did in the previous sections, we will also make this more explicit with the two examples of Haar and Mexican hat wavelets:

Haar wavelet. We consider the Haar wavelet defined as

$$\psi(x) \;=\; \begin{cases} -1\;, & x \in [-1,0) \\ 1\;, & x \in [0,1] \\ 0\;, & \text{else} \end{cases} \tag{51}$$

as negative first derivative of a hat function

$$h(x) \;=\; \begin{cases} 1-|x|\;, & x \le 1 \\ 0\;, & \text{else}\;\;. \end{cases} \tag{52}$$

A simple calculation shows that h can be written in terms of the box function χ considered below, namely $h = \chi * \chi$. Together we have $\psi = -\frac{d}{dx}(\chi * \chi)$. With (45) this implies in the Fourier domain

$$\tilde{\psi}(\xi) \;=\; -\frac{i\xi}{2\pi}\,\operatorname{sinc}^2\left(\frac{1}{2}\,\xi\right) \tag{53}$$

Written as pseudodifferential operator the basic convolution of a Haar wavelet transform looks as follows:

$$\psi * f \;=\; -\frac{1}{\sqrt{2\pi}}\,\frac{d}{dx}\,\operatorname{sinc}^2\left(\frac{1}{2i}\,\frac{d}{dx}\right) f \;\;. \tag{54}$$

For the scaled version of ψ we can use (50). The power series looks as follows:

$$\tilde{\psi}_\sigma * f \;=\; \sqrt{\frac{\sigma}{2\pi}}\,\frac{d}{dx}\,\mathrm{sinc}^2\left(\frac{-\sigma}{2i}\,\frac{d}{dx}\right) \;. \tag{55}$$

With this equation we have expressed the wavelet transform. The convolutions appearing in the back-transform are the same except the mirroring of the kernel. This operator can be used as an example for the linear transform L in the variational methods or evolution equations shown above.

Mexican hat wavelet. The Mexican hat mother wavelet is given by the second derivative of a Gaussian function $\psi(x) = \frac{d^2}{dx^2}\exp\left(-\frac{x^2}{2}\right)$. The Gaussian has already been considered in (47), so that we can directly give the corresponding derivative operator

$$\psi * f \;=\; \sqrt{2\pi}\,\frac{d^2}{dx^2}\,\exp\left(\frac{1}{2}\,\frac{d^2}{dx^2}\right) f \;=\; \sqrt{2\pi}\sum_{k=0}^{\infty}\frac{1}{2^k k!}\,\frac{d^{2k+2}}{dx^{2k+2}}\,f \;. \tag{56}$$

Again we consider scaled and mirrored versions since they appear as convolution kernels in the wavelet transform:

$$\tilde{\psi}_\sigma * f \;=\; -\sqrt{2\pi\sigma}\,\frac{d^2}{dx^2}\,\exp\left(\frac{\sigma^2}{2}\,\frac{d^2}{dx^2}\right) f \;=\; -\sqrt{2\pi\sigma}\sum_{k=0}^{\infty}\frac{\sigma^{2k}}{2^k k!}\,\frac{d^{2k+2}}{dx^{2k+2}}\,f \;. \tag{57}$$

These examples show that it is possible to reformulate the convolutions in the integrodifferential evolution equation (30) as power series of differential operators. This eliminates all integrations in this equation except of one: The outer integral over all scales σ still remains.

8. Conclusions

The goal of the present paper was to shed some light on the differences between continuous multiscale wavelet shrinkage on one hand and nonlinear diffusion filters of arbitrary order and their variational counterparts on the other hand. This has been achieved by deriving novel integrodifferential equations from multiscale wavelet shrinkage. To this end we exploited the fact that wavelets with a finite number of nonvanishing moments represent smoothed derivative operators. Our investigations apply to a broad class of widely-used wavelet types such as the Haar wavelet and the Mexican hat wavelet. The resulting integrodifferential equations differ from their nonlinear diffusion counterparts by the additional presmoothing of derivatives and integration over a continuum of scales. Moreover, they can be rewritten as a weighted average of pseudodifferential equations. It is our hope that our analysis will form a first step towards a more detailed quantitative characterisation between multiscale and iterative concepts for discontinuity-preserving denoising. We are currently extending our results to discrete considerations where very efficient multi-scale implementations are available. Moreover, we are investigating the algorithmic use of multiscale ideas within diffusion filtering, e.g. by truncated representations of the series of differential operators.

Acknowledgements

We gratefully acknowledge partly funding by the *Deutsche Forschungsgemeinschaft (DFG)*, project WE 2602/2-2. Moreover, we thank Gabriele Steidl (University of Mannheim) and Martin Welk (Saarland University) for helpful discussions on wavelet shrinkage.

References

[1] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. *Minimizing total variation flow*, Differential and Integral Equations, 14(2001) 321–360.

[2] Y. Bao and H. Krim. *Towards bridging scale-space and multiscale frame analyses*, In A. A. Petrosian and F. G. Meyer, editors, *Wavelets in Signal and Image Analysis*, volume 19 of Computational Imaging and Vision, chapter 6. Kluwer, Dordrecht, 2001.

[3] K. Bredies, D. A. Lorenz, and P. Maass. *Mathematical concepts of multiscale smoothing*, Applied and Computational Harmonic Analysis, 19(2005), 141–161.

[4] E. J. Candés and F. Guo. *New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction*, Signal Processing, 82(2002), 1519–1543.

[5] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM Journal on Numerical Analysis, 29(1992), 182–193.

[6] A. Chambolle, R. DeVore, N.-Y. Lee, and B. J. Lucier. *Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Transactions on Image Processing, 7(1998), 319–335.

[7] A. Chambolle and B. L. Lucier. *Interpreting translation-invariant wavelet shrinkage as a new image smoothing scale space*, IEEE Transactions on Image Processing, 10(2001), 993–1000.

[8] T. F. Chan and H. M. Zhou. *Total variation improved wavelet thresholding in image compression*, In "Proc. Seventh International Conference on Image Processing," volume II, pages 391–394, Vancouver, Canada 2000.

[9] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. *Two deterministic half-quadratic regularization algorithms for computed imaging*, Proc. IEEE International Conference on Image Processing (ICIP-94, Austin, Nov. 13-16, 1994), 2(1994), 168–172.

[10] A. Cohen, W. Dahmen, I. Daubechies, and R. DeVore. *Harmonic analysis in the space BV*, Revista Matematica Iberoamericana, 19(2003), 235–262.

[11] A. Cohen, R. DeVore, P. Petrushev, and H. Xu. *Nonlinear approximation and the space $BV(R^2)$*, American Journal of Mathematics, 121(1999), 587–628.

[12] R. R. Coifman and A. Sowa. *Combining the calculus of variations and wavelets for image enhancement*, Applied and Computational Harmonic Analysis, 9(2000), 1–18.

[13] R. R. Coifman and A. Sowa. *New methods of controlled total variation reduction for digital functions*, SIAM Journal on Numerical Analysis, 39(2001), 480–498.

[14] I. Daubechies. "Ten Lectures on Wavelets," SIAM, Philadelphia, 1992.

[15] I. Daubechies and G. Teschke. *Variational image restoration by means of wavelets: Simultaneous decomposition, deblurring and denoising*, Applied and Computational Harmonic Analysis, 19(2005), 1–16.

[16] S. Didas, J. Weickert, and B. Burgeth. *Stability and local feature enhancement of higher order nonlinear diffusion filtering*, in "Pattern Recognition"(eds. W. Kropatsch, R. Sablatnig, and A. Hanbury), volume 3663 of Lecture Notes in Computer Science, pages 451–458. Springer, 2005.

[17] D. L. Donoho. *De-noising by soft thresholding*, IEEE Transactions on Information Theory, 41(1995), 613–627.

[18] D. L. Donoho and I. M. Johnstone. *Ideal spatial adaption by wavelet shrinkage*, Biometrica, 81(1994), 425–455.

[19] R. Duits, L. Florack, J. de Graaf, and B. Ter Haar Romeny. *On the axioms of scale space theory*, Journal of Mathematical Imaging and Vision, 20(2004), 267–298.

[20] S. Durand and J. Froment. *Reconstruction of wavelet coefficients using total-variation minimization*, SIAM Journal on Scientific Computing, 24(2003), 1754–1767.

[21] S. Durand and M. Nikolova. *Restoration of wavelet coefficients by minimizing a specially designed objective function*, in "Proc. Second IEEE Workshop on Geometric and Level Set Methods in Computer Vision" (eds. O. Faugeras and N. Paragios), Nice, France, INRIA, 2003.

[22] H.-Y. Gao. *Wavelet shrinkage denoising using the non-negative garrote*, Journal of Computational and Graphical Statistics, 7(1998), 469–488.

[23] T. Iijima. *Basic theory on normalization of pattern (in case of typical one-dimensional pattern)*(in Japanese), Bulletin of the Electrotechnical Laboratory, 26(1962), 368–388.

[24] S. L. Keeling and R. Stollberger. *Nonlinear anisotropic diffusion filtering for multiscale edge enhancement*, Inverse Problems, 18(2002), 175–190.

[25] M. Lysaker, A. Lundervold, and X.-C. Tai. *Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time*, IEEE Transactions on Image Processing, 12(2003), 1579–1590.

[26] F. Malgouyres. *Mathematical analysis of a model which combines total variation and wavelet for image restoration*, Inverse Problems, 2(2002), 1–10.

[27] S. Mallat. "A Wavelet Tour of Signal Processing," Academic Press, San Diego, second edition, 1999.

[28] Y. Meyer. "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations," volume 22 of University Lecture Series, AMS, Providence, 2001.

[29] P. Mrázek and J. Weickert. Rotationally invariant wavelet shrinkage. In B. Michaelis and G. Krell, editors, "Pattern Recognition," volume 2781 of Lecture Notes in Computer Science , pages 156–163. Springer, Berlin, 2003.

[30] P. Mrázek, J. Weickert, and G. Steidl. *Diffusion-inspired shrinkage functions and stability results for wavelet denoising*, International Journal of Computer Vision, 64(2005), 171–186.

[31] P. Mrázek, J. Weickert, G. Steidl, and M. Welk. *On iterations and scales of nonlinear filters*, in "Proc. of the Computer Vision Winter Workshop 2003" (ed. O. Drbohlav), pages 61–66. Czech Pattern Recognition Society, 2003.

[32] M. Nielsen, L. Florack, and R. Deriche. *Regularization, scale-space and edge detection filters*, Journal of Mathematical Imaging and Vision, 7(1997), 291–307.

[33] W. J. Niessen, B. M. ter Haar Romeny, L. M. Florack, and M. A. Viergever. *A general framework for geometry-driven evolution equations*, International Journal of Computer Vision, 21(1997), 187–205.

[34] N. Nordström. *Biased anisotropic diffusion – a unified regularization and diffusion approach to edge detection,* Image and Vision Computing, 8(1990), 318–327.

[35] P. Perona and J. Malik. *Scale space and edge detection using anisotropic diffusion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1990), 629–639.

[36] L. I. Rudin, S. Osher, and E. Fatemi. *Nonlinear total variation based noise removal algorithms*, Physica D, 60(1992), 259–268.

[37] O. Scherzer. *Denoising with higher order derivatives of bounded variation and an application to parameter estimation*, Computing, 60(1998), 1–27.

[38] O. Scherzer and J. Weickert. *Relations between regularization and diffusion filtering*, Journal of Mathematical Imaging and Vision, 12(2000), 43–63.

[39] C. Schnörr. *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, Journal of Mathematical Imaging and Vision, 4(1994)189–198.

[40] J.-L. Starck, M. Elad, and D. L. Donoho. *Image decomposition via the combination of sparse representations and a variational approach*, IEEE Transactions on Image Processing, 14(2005):1570–1582.

[41] 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*, SIAM Journal on Numerical Analysis , 42(2004), 686–713.

[42] M. E. Taylor. "Pseudodifferential Operators," Princeton University Press, Princeton, New Jersey, 1981.

[43] M. E. Taylor. "Partial Differential Equations I – Basic Theory," Springer, New York, 1996.

[44] J. Weickert. "Anisotropic Diffusion in Image Processing," B. G. Teubner, Stuttgart, 1998.

[45] J. Weickert, G. Steidl, P. Mrázek, M. Welk, and T. Brox. *Diffusion filters and wavelets: What can they learn from each other?*, in "Handbook of Mathematical Models in Computer Vision" (eds. N. Paragios, Y. Chen, and O. Faugeras). Springer, New York, 2006.

[46] M. Welk, G. Steidl, and J. Weickert. *Locally analytic schemes: A link between diffusion filtering and wavelet shrinkage*, Technical Report 2100, Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN, U.S.A., 2006.

[47] D. Werner. "Funktional Analysis," Springer, Berlin, third edition, 2000.

[48] A. P. Witkin. *Scale-space filtering*, In "Proceedings of the 8th International Joint Conference on Artificial Intelligence" volume 2, pages 945–951, Karlsruhe, Germany, August 1983.

[49] Y.-L. You and M. Kaveh. *Fourth-order partial differential equations for noise removal*, IEEE Transactions on Image Processing, 9(2000), 1723–1730.

Received and revised September 2006.

E-mail address: didas@mia.uni-saarland.de
E-mail address: weickert@mia.uni-saarland.de