From Two-Dimensional Nonlinear Diffusion to Coupled Haar Wavelet Shrinkage

Pavel Mrázek[1] and Joachim Weickert[2]

[1] UPEK, Husinecká 7, 13000 Praha 3, Czech Republic
pavel.mrazek@upek.com
http://www.upek.com

[2] Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science, Building 27.1
Saarland University, 66123 Saarbrücken, Germany
weickert@mia.uni-saarland.de
http://www.mia.uni-saarland.de

August 9, 2005

Abstract

This paper studies the connections between discrete two-dimensional schemes for shift-invariant Haar wavelet shrinkage on one hand, and nonlinear diffusion on the other. We show that using a single iteration on a single scale, the two methods can be made equivalent by the choice of the nonlinearity which controls each method: the shrinkage function, or the diffusivity function, respectively. In the two-dimensional setting, this diffusion–wavelet connection shows an important novelty compared to the one-dimensional framework or compared to classical 2-D wavelet shrinkage: The structure of two-dimensional diffusion filters suggests to use a coupled, synchronised shrinkage of the individual wavelet coefficient channels. This coupling enables to design Haar wavelet filters with good rotation invariance at a low computational cost. Furthermore, by transferring the channel coupling of vector- and matrix-valued nonlinear diffusion filters to the Haar wavelet setting, we obtain well-synchronised shrinkage methods for colour and tensor images. Our experiments show that these filters perform significantly better than conventional shrinkage methods that process all wavelets independently.

1 Introduction

Wavelet shrinkage and nonlinear diffusion are two seemingly very different concepts for discontinuity-preserving signal and image denoising. Since they are serving the same purpose, however, it would be desirable to understand if there are intrinsic connections between both worlds. On one hand this may allow to transfer results from one framework to the other, on the other hand it may allow to design hybrid method that combine advantages from both concepts.

Although research in this direction is still a relatively young field, already a number of interesting connections between wavelet shrinkage, partial differential equations (PDEs) and related regularisation methods has been established. Most of them analyse the *continuous* framework [4, 5, 8, 7, 28, 35, 36] or focus on designing methods that use wavelet shrinkage and PDE-based denoising methods in combination [3, 6, 10, 14, 25, 26].

Regarding the relations between wavelet shrinkage of *discrete* signals and PDE-based denosing, not much research has been done so far. One notable exception is a recent paper by Coifman and Sowa [11] where they propose total variation (TV) diminishing flows that act along the direction of Haar wavelets. Bao and Krim [2] addressed the problem of texture loss in diffusion scale-spaces by incorporating ideas from wavelet analysis. An experimental evaluation of the denoising capabilities of 3-D wavelet shrinkage and nonlinear diffusion filters is presented in a paper by Frangakis et al. [17].

Also in our recent work we have analysed discrete relations between nonlinear diffusion and wavelet shrinkage in the *one-dimensional* setting. By deriving identical analytical solutions we proved equivalence between shift invariant soft Haar wavelet shrinkage on a single scale, space-discrete (but time-continuous) nonlinear diffusion with a TV diffusivity, and discrete TV regularisation [39]. Using these ideas on multiple scales and iterating the method comes down to hybrid techniques that aim to combine the efficiency of wavelets with the quality of PDE-based methods. Their performance is evaluated in [32]. Considering space-discrete nonlinear diffusion and replacing the time-continuous formulation by an explicit (Euler forward) time discretisation allowed us to find general relations between the diffusivity for nonlinear diffusion filtering and the shrinkage function of shift invariant Haar wavelet shrinkage on a single scale [30]. In this way we identified also nonlinear diffusivities for hard, firm and Garrote wavelet shrinkage, and we proposed novel shrinkage functions that were inspired from nonlinear diffusivities and offered competitive performance. Moreover, this connection enabled us to derive novel stability results for single-scale wavelet shrinkage, including monotonicity preservation and sign stability [31].

Since the connection between fully discrete nonlinear diffusion filtering and wavelet shrinkage turned out to be very fruitful in the *one-dimensional* setting, it is natural to ask whether we can also gain novel insights in the *two-dimensional* case by exploiting a similar strategy. This is the topic of the present paper where we focus mainly on one important novelty compared to 1-D filtering, namely *channel coupling* and its consequences. The link to nonlinear diffusion requires that the shrinkage function of individual channels of wavelet coefficients is synchronised via a joint estimation of the image gradient. The resulting coupled wavelet shrinkage then inherits some desirable properties from the corresponding nonlinear diffusion filter, namely a good approximation of rotation invariance and a straightforward extension to filtering of vector- and tensor-valued images.

The property of rotation invariance is natural for diffusion, being inherent in the continuous formulation of the diffusion process. Unfortunately, this quality is usually missing in wavelet shrinkage methods. The relation between the two techniques allows to construct rotation-invariant wavelet shrinkage methods easily, with very little additional computational cost.

Diffusion methods have an established approach to filtering of vector-, or even tensor-valued data. Through a common estimation of diffusivity, the individual data components evolve in a synchronised manner, and possible artifacts or misalignment of features in the individual data channels are avoided. Having established the link to coupled wavelet shrinkage, wavelet-based methods suitable for filtering of vector- and tensor-valued images can be designed in a straightforward way.

We have presented first results on diffusion-inspired rotationally invariant wavelet shrinkage in a recent conference paper [29]. The present paper gives substantial additional results: It exploits alternative numerical schemes and analyses them in detail. These schemes offer additional degrees of freedom that can be used to achieve better approximations to the rotationally invariant case. Moreover, the results on diffusion-inspired coupled wavelet shrinkage for colour and tensor images are described here for the first time.

Our paper is organised as follows. Section 2 provides a brief introduction to Haar wavelet shrinkage in one and two dimensions. Section 3 introduces nonlinear diffusion and its explicit discretisation. Three different discrete diffusion schemes are then related to wavelet shrinkage in Section 4. The rotation invariance of the resulting coupled wavelet shrinkage is studied in Section 5, and extended to colour images in Section 6. Coupled wavelet shrinkage for tensor-valued data is investigated in Section 7. We conclude the paper with a summary in Section 8.

2 Wavelet Shrinkage

2.1 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

$$f = \sum_{i\in\mathbb{Z}} \langle f, \varphi_i^n \rangle \, \varphi_i^n + \sum_{j=-\infty}^{n} \sum_{i\in\mathbb{Z}} \langle f, \psi_i^j \rangle \, \psi_i^j, \tag{1}$$

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 [27]. 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 (1).

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:

$$u := \sum_{i\in\mathbb{Z}} \langle f, \varphi_i^n \rangle \, \varphi_i^n + \sum_{j=-\infty}^{n} \sum_{i\in\mathbb{Z}} S_\theta(\langle f, \psi_i^j \rangle) \, \psi_i^j. \tag{2}$$

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

$$\begin{aligned} \psi(x) &= \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}, & (3)\\ \phi(x) &= \mathbf{1}_{[0,1)} & (4) \end{aligned}$$

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:

$$c_i^j = \frac{c_{2i}^{j-1} + c_{2i+1}^{j-1}}{\sqrt{2}}, \qquad d_i^j = \frac{c_{2i}^{j-1} - c_{2i+1}^{j-1}}{\sqrt{2}}, \tag{5}$$

and

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

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 the literature [12, 18, 19, 27]. In the present paper we employ hard shrinkage for the experiments, since it performs particularly well w.r.t. image denoising [30]:

$$S_\theta(x) = \begin{cases} 0 & \text{for } |x| \le \theta, \\ x & \text{for } |x| > \theta. \end{cases} \tag{7}$$

2.2 Discrete Translation-Invariant Scheme in 1-D

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 (5), shrinkage, and synthesis (6). 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 (5) and (6), Haar wavelet shrinkage on one level produces the signal $\mathbf{u}^+ = (u_i^+)_{i=0}^{N-1}$ with coefficients

$$\begin{aligned} u_{2i}^+ &= \frac{c_i + S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right), & (8)\\ u_{2i+1}^+ &= \frac{c_i - S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right). & (9) \end{aligned}$$

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 [9] 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* [21]. 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

$$u_{2i-1}^- = \frac{f_{2i-1}+f_{2i}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1}-f_{2i}}{\sqrt{2}}\right), \tag{10}$$

$$u_{2i}^- = \frac{f_{2i-1}+f_{2i}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1}-f_{2i}}{\sqrt{2}}\right). \tag{11}$$

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

$$\begin{aligned} u_i &= \frac{u_i^- + u_i^+}{2} \\ &= \frac{f_{i-1}+2f_i+f_{i+1}}{4} + \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_i - f_{i+1}}{\sqrt{2}}\right) - \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_{i-1}-f_i}{\sqrt{2}}\right) \end{aligned} \tag{12}$$

2.3 Shrinkage in Two Dimensions

The easiest way to design a two-dimensional wavelet transform is to use separable filters [40]. The 2-D wavelet transform then describes a 2-D signal $\mathbf{f} = (f_{i,j})$ with $i = 0,...,N_x-1$ and $j = 0,...,N_y-1$ by its low-pass component at level n, $\mathbf{v}^n$, and three channels of wavelet coefficients $\mathbf{w}_x^l$, $\mathbf{w}_y^l$ and $\mathbf{w}_{xy}^l$ at levels $l = 1,\ldots,n$. This wavelet representation is created by an alternating application of the one-dimensional low-pass and high-pass filters L and H in the directions of the axes x and y:

$$\mathbf{v}^{l+1} = L(x) * L(y) * \mathbf{v}^l, \tag{13}$$

$$\mathbf{w}_y^{l+1} = L(x) * H(y) * \mathbf{v}^l, \tag{14}$$

$$\mathbf{w}_x^{l+1} = H(x) * L(y) * \mathbf{v}^l, \tag{15}$$

$$\mathbf{w}_{xy}^{l+1} = H(x) * H(y) * \mathbf{v}^l, \tag{16}$$

with the initial condition $\mathbf{v}^0 = \mathbf{f}$. For image smoothing, the wavelet coefficients $\mathbf{w}_x, \mathbf{w}_y, \mathbf{w}_{xy}$ are subjected to a shrinkage function S, and the filtered image $\mathbf{u}$ is reconstructed from the shrunken coefficients using an inverse procedure to (13)–(16). The Haar wavelet transform is described by a low-pass filter L with coefficients $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, and a high-pass filter H with coefficients $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ [40].

Let us now consider a *single* decomposition level, and the wavelet shrinkage steps which contribute to the output pixel $u_{i,j}$. Using the translation-invariant scheme [9], we have to consider the four 2×2 neighbourhoods in which the pixel (i,j) is involved. We denote by upper index α the neighbourhood $\{i,i+1\} \times \{j,j+1\}$; by β the positions $\{i,i+1\} \times \{j-1,j\}$; by γ the neighbourhood $\{i-1,i\} \times \{j,j+1\}$; and, finally, by δ the positions $\{i-1,i\} \times \{j-1,j\}$.

$$v^\alpha = \begin{array}{|c|c|c|}\hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_y^\alpha = \begin{array}{|c|c|c|}\hline 0 & -\frac{1}{2} & -\frac{1}{2} \\ \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_x^\alpha = \begin{array}{|c|c|c|}\hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_{xy}^\alpha = \begin{array}{|c|c|c|}\hline 0 & -\frac{1}{2} & \frac{1}{2} \\ \hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

$$v^\beta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline \end{array} \quad w_y^\beta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & -\frac{1}{2} & -\frac{1}{2} \\ \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline \end{array} \quad w_x^\beta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline \end{array} \quad w_{xy}^\beta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & -\frac{1}{2} & \frac{1}{2} \\ \hline 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline \end{array}$$

$$v^\gamma = \begin{array}{|c|c|c|}\hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_y^\gamma = \begin{array}{|c|c|c|}\hline -\frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_x^\gamma = \begin{array}{|c|c|c|}\hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad w_{xy}^\gamma = \begin{array}{|c|c|c|}\hline -\frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

$$v^\delta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \end{array} \quad w_y^\delta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline -\frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \end{array} \quad w_x^\delta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \end{array} \quad w_{xy}^\delta = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline -\frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & -\frac{1}{2} & 0 \\ \hline \end{array}$$

Figure 1: The first-level Haar wavelet coefficients expressed using 3×3 masks centered at the pixel i, j. The masks represent multiplication of the input signal with the given coefficients, so e.g. $w_x^\alpha = \frac{1}{2} f_{i,j} + \frac{1}{2} f_{i,j+1} - \frac{1}{2} f_{i+1,j} - \frac{1}{2} f_{i+1,j+1}$.

The input signal in neighbourhood α is first transformed into v^α, w_y^α, w_x^α and w_{xy}^α; see Fig. 1 for the definition of the corresponding masks. The wavelet coefficients $w_y^\alpha, w_x^\alpha, w_{xy}^\alpha$ are then subjected to a shrinkage function S, and the (i, j) pixel of the filtered signal belonging to the neighbourhood α is obtained using

$$u_{i,j}^\alpha \;=\; \frac{1}{2} \left(v^\alpha + S(w_x^\alpha) + S(w_y^\alpha) + S(w_{xy}^\alpha) \right). \tag{17}$$

Similar expressions can be derived for the results arising from the neighbourhoods β, γ and δ; the necessary masks are shown in Fig. 1. To obtain the final result of a shift-invariant 2-D Haar wavelet shrinkage on a single level, the four intermediate results $u_{i,j}^\alpha$, $u_{i,j}^\beta$, $u_{i,j}^\gamma$ and $u_{i,j}^\delta$ have to be averaged. The complete formula for a single-level Haar wavelet shrinkage filter then reads

$$\begin{aligned} u_{i,j} \;=\; & \frac{1}{8} \left(v^\alpha + S(w_x^\alpha) + S(w_y^\alpha) + S(w_{xy}^\alpha) \right) \\ + \; & \frac{1}{8} \left(v^\beta + S(w_x^\beta) - S(w_y^\beta) - S(w_{xy}^\beta) \right) \\ + \; & \frac{1}{8} \left(v^\gamma - S(w_x^\gamma) + S(w_y^\gamma) - S(w_{xy}^\gamma) \right) \\ + \; & \frac{1}{8} \left(v^\delta - S(w_x^\delta) - S(w_y^\delta) + S(w_{xy}^\delta) \right). \end{aligned} \tag{18}$$

This formula will be used for establishing correspondences to nonlinear diffusion filtering.

3 Nonlinear Diffusion

3.1 Basic Concept

The basic idea behind nonlinear diffusion filtering [33] is to obtain a family $u(x, y, t)$ of filtered versions of the signal $f(x, y)$ as the solution of a suitable diffusion process

$$u_t = \operatorname{div}\left(g(|\nabla u|^2)\,\nabla u\right) \tag{19}$$

with f as initial condition: $u(x,y,0) = f(x,y)$. Here subscripts denote partial derivatives, and the diffusion time t is a simplification parameter: larger values correspond to more pronounced filtering.

The diffusivity $g(|\nabla u|^2)$ is a nonnegative function that controls the amount of diffusion. Usually, it is decreasing in $|\nabla u|^2$. This ensures that strong edges are less blurred by the diffusion filter than noise and low-contrast details. A typical representative of a nonlinear diffusivity is given by [33]

$$g\big(s^2\big) = \frac{1}{1+\frac{s^2}{\lambda^2}}. \tag{20}$$

3.2 Explicit Discretisation Scheme

When applied to discrete signals, the partial differential equation (19) has to be discretised. A filtered solution is then found by an iterative procedure, starting from the noisy signal at time 0, $\mathbf{u}^0 = \mathbf{f}$, and proceeding by $\mathbf{u}^{k+1} = F(\mathbf{u}^k)$, $k = 0, 1, 2, \ldots$. In this paper, we focus on explicit finite difference schemes which in each iteration apply simple operations to neighbouring pixels.

The divergence expression on the right hand side of (19) can be decomposed in 2-D by means of two orthonormal basis vectors $\mathbf{v_1}$ and $\mathbf{v_2}$:

$$\operatorname{div}\left(g(|\nabla u|^2)\,\nabla u\right) = \sum_{p=1}^{2} \partial_{\mathbf{v_p}}\left(g(|\nabla u|^2)\,\partial_{\mathbf{v_p}} u\right). \tag{21}$$

Replacing the derivatives in (19), (21) by finite differences, we can write the explicit finite difference discretisation of the nonlinear diffusion as

$$u_{i,j}^{k+1} = u_{i,j}^k + \tau \sum_{(m,n)\in\mathcal{N}(i,j)} g_{m,n}^k \,\frac{u_{m,n}^k - u_{i,j}^k}{|(m,n)-(i,j)|^2}. \tag{22}$$

Here the upper index k denotes solution at time $k\tau$ with τ standing for the time step, the set $\mathcal{N}(i,j)$ contains the neighbours of pixel (i,j), and the expression $|(i,j)-(m,n)|$ stands for the distance between pixels (i,j) and (m,n). The term $g_{m,n}^k$ approximates $g(|\nabla u(x,y,t)|^2)$ at location $(\frac{i+m}{2}, \frac{j+n}{2})$ and time $k\tau$. It represents the diffusivity belonging to the connection between the (i,j) and (m,n), where the gradient magnitude $|\nabla u(x,y,t)|^2$ can be estimated from discrete data using 2×2 masks. This topic is addressed in detail in Section 4.4.

The neighbourhood connectivity between pixels depends on the choice of basis vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ for divergence estimation in (21); these vectors influence which pixels are included in the neighbourhood set $\mathcal{N}(i,j)$. In the following we consider three types of connectivity: diagonal, vertical/horizontal (4-neighbourhood), and combined (8-neighbourhood). Each of them leads to a slightly different discrete scheme for the nonlinear diffusion equation, and each can be related to wavelet shrinkage. These relations are studied next.

4 Diffusion-Wavelet Connections

4.1 Diagonal Diffusion Connectivity

Choosing the basis vectors for divergence discretisation (21) in diagonal directions as $\mathbf{v_1} := (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and $\mathbf{v_2} := (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$, the explicit nonlinear diffusion step (22) becomes

$$u_{i,j}^{k+1} = u_{i,j}^k + \tau \sum_{(m,n)\in\mathcal{D}(i,j)} g_{m,n}^k \, \frac{u_{m,n}^k - u_{i,j}^k}{2}. \tag{23}$$

Here, the set $\mathcal{D}(i,j)$ contains the diagonal neighbours of pixel (i,j), and the grid size is assumed to be 1.

The diagonal discretisation of nonlinear diffusion given by (23) contains no vertical or horizontal connections between neighbouring pixels (see Fig. 2 left). This scheme has the drawback of decoupling the image into two overlaying diagonal grids which are connected only at the image boundaries and may create some checkerboard-like structures during the image evolution. However, this diagonal discretisation represents a consistent finite difference approximation to the continuous equation. It has been used successfully by Keeling and Stollberger [22]. Since its spatial consistency can be shown to be of second order, the rotation invariance of the continuous equation is approximated well.

Let us now investigate the connection between a single-level wavelet shrinkage (18) and an explicit diffusion iteration (23). To this end, we consider the first diffusion iteration, starting from the initial signal $\mathbf{f} = (f_{i,j})$ and creating a solution $\mathbf{u} = (u_{i,j})$. We will express the diffusion iteration in the terms of the wavelet coefficients from Section 2.

For the first iteration, (23) becomes

$$u_{i,j} = f_{i,j} + \tau \sum_{(m,n)\in\mathcal{D}(i,j)} g_{m,n} \, \frac{f_{m,n} - f_{i,j}}{2}. \tag{24}$$

The first term on the right-hand side can be rewritten as

$$\begin{aligned} f_{i,j} &= \frac{1}{8} \begin{array}{|c|c|c|}\hline \frac{1}{2} & 1 & \frac{1}{2} \\ \hline 1 & 2 & 1 \\ \hline \frac{1}{2} & 1 & \frac{1}{2} \\ \hline \end{array} + \frac{1}{8} \begin{array}{|c|c|c|}\hline -\frac{1}{2} & -1 & -\frac{1}{2} \\ \hline -1 & 6 & -1 \\ \hline -\frac{1}{2} & -1 & -\frac{1}{2} \\ \hline \end{array} \\ &= \frac{1}{8}\left(v^\alpha + v^\beta + v^\gamma + v^\delta\right) \\ &+ \frac{1}{8}\left(w_x^\alpha + w_y^\alpha + w_{xy}^\alpha + w_x^\beta - w_y^\beta - w_{xy}^\beta\right. \\ &\quad \left. -w_x^\gamma + w_y^\gamma - w_{xy}^\gamma - w_x^\delta - w_y^\delta + w_{xy}^\delta\right) \end{aligned} \tag{25}$$

where the 3×3 boxes stand for a mask multiplication with the input signal, and the v and w represent the wavelet coefficients for position (i,j); see Fig. 1. Then, the gradient magnitude for the diffusivity calculation has to be estimated from the discrete samples. In Section 4.4 we will show that the diffusivity value may be obtained as a function of the wavelet coefficients w_x, w_y and w_{xy} such that it can be written as

$$g_{m,n} = g^\omega := g\big((w_x^\omega)^2 + (w_y^\omega)^2 + c \cdot (w_{xy}^\omega)^2\big) \tag{26}$$

where

$$\omega := \begin{cases} \alpha & \text{if } (m,n) = (i+1, j+1) \\ \beta & \text{if } (m,n) = (i+1, j-1) \\ \gamma & \text{if } (m,n) = (i-1, j+1) \\ \delta & \text{if } (m,n) = (i-1, j-1) \end{cases} \tag{27}$$

and c is an arbitrary nonnegative constant.
Finally, the last term from (24) may be expressed using wavelet coefficients as

$$f_{m,n} - f_{i,j} = \begin{cases} -w_x^\alpha - w_y^\alpha & \text{if } (m,n) = (i+1, j+1) \\ -w_x^\beta + w_y^\beta & \text{if } (m,n) = (i+1, j-1) \\ w_x^\gamma - w_y^\gamma & \text{if } (m,n) = (i-1, j+1) \\ w_x^\delta + w_y^\delta & \text{if } (m,n) = (i-1, j-1). \end{cases} \tag{28}$$

To summarise (24)–(28), we can write a single iteration of nonlinear diffusion using the wavelet decomposition components v and w in the form

$$\begin{aligned} u_{i,j} &= \frac{1}{8} \left(v^\alpha + w_x^\alpha(1 - 4\tau g^\alpha) + w_y^\alpha(1 - 4\tau g^\alpha) + w_{xy}^\alpha\right) \\ &+ \frac{1}{8} \left(v^\beta + w_x^\beta(1 - 4\tau g^\beta) - w_y^\beta(1 - 4\tau g^\beta) - w_{xy}^\beta\right) \\ &+ \frac{1}{8} \left(v^\gamma - w_x^\gamma(1 - 4\tau g^\gamma) + w_y^\gamma(1 - 4\tau g^\gamma) - w_{xy}^\gamma\right) \\ &+ \frac{1}{8} \left(v^\delta - w_x^\delta(1 - 4\tau g^\delta) - w_y^\delta(1 - 4\tau g^\delta) + w_{xy}^\delta\right). \end{aligned} \tag{29}$$

Comparing the diffusion iteration (29) and the single-level wavelet shrinkage (18), we observe that the two equations are equivalent under the conditions

$$S(w_x^\omega) = w_x^\omega \, (1 - 4\,\tau\, g^\omega)\,, \tag{30}$$

$$S(w_y^\omega) = w_y^\omega \, (1 - 4\,\tau\, g^\omega)\,, \tag{31}$$

$$S(w_{xy}^\omega) = w_{xy}^\omega. \tag{32}$$

Equations (30)–(32) connect the diffusivity function g controlling nonlinear diffusion (discretised with diagonal neighbourhood connectivity) to the shrinkage function S of wavelet shrinkage. If these conditions hold true, both two-dimensional procedures (limited to a single scale / single iteration) are equivalent.

The equations (30), (31) are similar to the one-dimensional situation which was analysed in detail in [30]. The surprising fact in the 2-D equations (30)–(32) is the use of different shrinkage rules for the different channels of wavelet coefficients, while the classical wavelet shrinkage applies the same shrinkage function S to each of them separately. In (30)–(31), the shrinkage of w_x and w_y is interconnected via the joint diffusivity term g^ω; the third channel, w_{xy}, is left unshrunken by (32).

4.2 Vertical/Horizontal Diffusion Connectivity

In this section we again formulate the explicit diffusion iteration using wavelet coefficients, but this time the vertical and horizontal connectivity between neighbouring pixels is employed. In this case, the first diffusion iteration can be written as

$$u_{i,j} = f_{i,j} + \tau \sum_{(m,n) \in \mathcal{V}(i,j)} g_{m,n} \, (f_{m,n} - f_{i,j}) \tag{33}$$

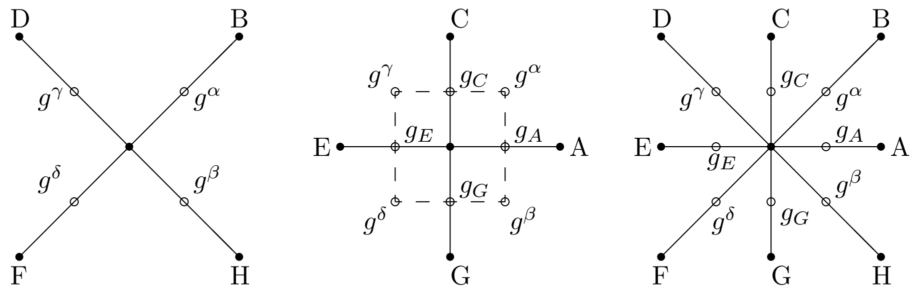

Figure 2: Geometry of the discrete diffusion process and the diffusivity estimation for the three studied schemes. **Left:** diagonal connectivity. **Middle:** horizontal/vertical connectivity. **Right:** combined scheme, full connectivity in the 8-neighbourhood.

where the set $\mathcal{V}(i,j)$ contains the 4-neighbours of the pixel (i,j): $\mathcal{V}(i,j) = \{(i+1,j),(i,j+1),(i-1,j),(i,j-1)\}$. The four neighbours will be denoted by A, C, E, G, respectively, as in Fig. 2 center. This numerical scheme with horizontal and vertical connectivity is obtained by discretising the divergence expression (21) using the standard basis vectors $\mathbf{v_1} := (1,0)$ and $\mathbf{v_2} := (0,1)$.

The first term $f_{i,j}$ in (33) can be rewritten using wavelet coefficients in the same way as before, see (25).

As for the diffusivity g, we need to obtain its value at the connection between the center pixel and its neighbours. We will estimate it from the average of the nearby diffusivities g^ω (see Fig. 2 center):

$$g_A = \frac{1}{2}(g^\alpha + g^\beta), \qquad g_E = \frac{1}{2}(g^\gamma + g^\delta), \tag{34}$$

$$g_C = \frac{1}{2}(g^\alpha + g^\gamma), \qquad g_G = \frac{1}{2}(g^\beta + g^\delta). \tag{35}$$

The last term in (33), $f_{m,n} - f_{i,j}$, can be expressed using wavelet coefficients in two ways for each of the pixels (m,n):

$$f_{m,n} - f_{i,j} = \begin{cases} d_A := -w_x^\alpha - w_{xy}^\alpha = -w_x^\beta + w_{xy}^\beta & \text{for } (m,n) = (i+1,j), \\ d_C := -w_y^\alpha - w_{xy}^\alpha = -w_y^\gamma + w_{xy}^\gamma & \text{for } (m,n) = (i,j+1), \\ d_E := w_x^\gamma + w_{xy}^\gamma = w_x^\delta - w_{xy}^\delta & \text{for } (m,n) = (i-1,j), \\ d_G := w_y^\beta + w_{xy}^\beta = w_y^\delta - w_{xy}^\delta & \text{for } (m,n) = (i,j-1). \end{cases} \tag{36}$$

To summarise, we can write the diffusion iteration (33) in terms of wavelet coefficients as

$$\begin{aligned} u_{i,j} &= \frac{1}{8}\left(v^\alpha + v^\beta + v^\gamma + v^\delta\right) \\ &+ \frac{1}{8}\big(w_x^\alpha + w_y^\alpha + w_{xy}^\alpha + w_x^\beta - w_y^\beta - w_{xy}^\beta \\ &\quad -w_x^\gamma + w_y^\gamma - w_{xy}^\gamma - w_x^\delta - w_y^\delta + w_{xy}^\delta\big) \\ &- \tau\frac{g^\alpha + g^\beta}{2}d_A - \tau\frac{g^\alpha + g^\gamma}{2}d_C - \tau\frac{g^\gamma + g^\delta}{2}d_E - \tau\frac{g^\beta + g^\delta}{2}d_G. \end{aligned} \tag{37}$$

Let us multiply the terms including d_X in the last line of (37) by the corresponding g^ω, and substitute for d_X from (36) so that the diffusivity g^ω multiplies only the corresponding coefficients w^ω. As an example, we have

$$g^\alpha d_A + g^\beta d_A \;=\; g^\alpha\big(-w_x^\alpha - w_{xy}^\alpha\big) + g^\beta\big(-w_x^\beta + w_{xy}^\beta\big).$$

Performing similar substitution for the other terms, and assembling all the instances of each wavelet coefficient together, we get

$$\begin{aligned} u_{i,j} \;=\; & \frac{1}{8}\,\big(v^\alpha + w_x^\alpha(1-4\tau g^\alpha) + w_y^\alpha(1-4\tau g^\alpha) + w_{xy}^\alpha(1-8\tau g^\alpha)\big) \\ +\; & \frac{1}{8}\,\big(v^\beta + w_x^\beta(1-4\tau g^\beta) - w_y^\beta(1-4\tau g^\beta) - w_{xy}^\beta(1-8\tau g^\beta)\big) \\ +\; & \frac{1}{8}\,\big(v^\gamma - w_x^\gamma(1-4\tau g^\gamma) + w_y^\gamma(1-4\tau g^\gamma) - w_{xy}^\gamma(1-8\tau g^\gamma)\big) \\ +\; & \frac{1}{8}\,\big(v^\delta - w_x^\delta(1-4\tau g^\delta) - w_y^\delta(1-4\tau g^\delta) + w_{xy}^\delta(1-8\tau g^\delta)\big)\,. \end{aligned} \tag{38}$$

Comparing (38) with (18) we see that a single-level wavelet shrinkage is identical to a single iteration of nonlinear diffusion discretised using the vertical/horizontal connectivity, under the condition that the diffusivity function g and the shrinkage function S satisfy

$$S(w_x^\omega) \;=\; w_x^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{39}$$
$$S(w_y^\omega) \;=\; w_y^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{40}$$
$$S(w_{xy}^\omega) \;=\; w_{xy}^\omega\,(1-8\,\tau\,g^\omega)\,. \tag{41}$$

We observe that – unlike in the scheme with diagonal connectivity – also the mixed wavelet coefficients w_{xy}^ω are shrunken now, and even at a faster rate than the coefficients w_x^ω, w_y^ω.

4.3 Combined Scheme

Both the scheme with diagonal connectivity (Sec. 4.1) and the one with vertical/horizontal connectivity (Sec. 4.2) represent numerically consistent discretisations of the nonlinear diffusion equation (19). They are of first order in time and second order in space. Therefore, any convex combination of these schemes is also a consistent approximation with the same consistency order. Such a combined scheme corresponds to a diffusion process that is discretised using the full pixel connectivity in the 8-neighbourhood as is illustrated in Fig. 2 right. A combined nonlinear diffusion scheme with weight $(1-q)$ for the diagonal and weight q for the vertical/horizontal connectivity corresponds to the shrinkage rules

$$S(w_x^\omega) \;=\; w_x^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{42}$$
$$S(w_y^\omega) \;=\; w_y^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{43}$$
$$S(w_{xy}^\omega) \;=\; w_{xy}^\omega\,(1-8\,q\,\tau\,g^\omega)\,. \tag{44}$$

In the special case when both schemes contribute with the same weight, we have $q=\frac{1}{2}$ and all the three wavelet channels are shrunken with identical speed:

$$S(w_x^\omega) \;=\; w_x^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{45}$$
$$S(w_y^\omega) \;=\; w_y^\omega\,(1-4\,\tau\,g^\omega)\,, \tag{46}$$
$$S(w_{xy}^\omega) \;=\; w_{xy}^\omega\,(1-4\,\tau\,g^\omega)\,. \tag{47}$$

This identical shrinkage for all channels resembles most the classical wavelet shrinkage approach, and seems to represent the best choice for noniterated filtering: all three coefficient channels w_x^ω, w_y^ω, w_{xy}^ω at uninteresting locations can be shrunken to zero in a single step.

It should be noted that the three channels are still coupled by the joint term g^ω on the right hand side of (45)–(47). As can be seen from (26), this joint diffusivity combines information from different wavelet channels, where the channel coupling originates from the diffusion equation (19). The question how to estimate the joint diffusivity g^ω from the discrete data samples is addressed next.

4.4 Diffusivity Estimation

The behaviour of a nonlinear diffusion filter (19) is controlled by the diffusivity function $g(|\nabla u|^2)$. Its value is obtained after estimating the gradient magnitude from the discrete data samples. There are several possibilities how to construct such an estimate, and how to express it using the available wavelet coefficients. This topic is addressed in the present section.

In all the diffusion discretisations presented above, the diffusivity value, and hence the (squared) gradient magnitude was needed in the center of four adjacent pixels (point O in Fig. 3). To estimate it, we will use a 2 by 2 neighbourhood only. In the following, the symbol ∂_x stands for the partial derivative with respect to x, ∂_y with respect to y. Let us now discuss different possibilities to approximate the squared gradient at point O.

Discretisation 1: Averaging before squaring. One way to estimate $|\nabla u|^2(O)$ is to first evaluate the partial derivatives at point O as an average of local derivative estimation at points P, R for the x-derivative, and at points S, Q for the y-derivative. Using this scheme, the gradient magnitude is calculated as

$$\begin{aligned}
|\nabla u|^2(O) &= \big(\partial_x u(O)\big)^2 + \big(\partial_y u(O)\big)^2 \\
&\approx \left(\frac{\partial_x u(P) + \partial_x u(R)}{2}\right)^2 + \left(\frac{\partial_y u(S) + \partial_y u(Q)}{2}\right)^2 \\
&\approx \begin{array}{|c|c|c|} \hline 0 & -\frac{1}{2} & \frac{1}{2} \\ \hline 0 & -\frac{1}{2} & \frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array}^{\,2} + \begin{array}{|c|c|c|} \hline 0 & \frac{1}{2} & \frac{1}{2} \\ \hline 0 & -\frac{1}{2} & -\frac{1}{2} \\ \hline 0 & 0 & 0 \\ \hline \end{array}^{\,2} \\
&= \big(-w_x^\alpha\big)^2 + \big(-w_y^\alpha\big)^2 \\
&= \big(w_x^\alpha\big)^2 + \big(w_y^\alpha\big)^2, \qquad (48)
\end{aligned}$$

where the 3×3 masks represent operations on the input signal as in Fig. 1. We see that the gradient magnitude can be estimated using the x and y channels of the wavelet coefficients. By means of Taylor expansions around O, one can verify that this estimate represents

$$|\nabla u|^2 + \frac{1}{12}h^2(u_x u_{xxx} + 3u_x u_{xyy} + 3u_y u_{xxy} + u_y u_{yyy}) + \mathcal{O}(h^4) \qquad (49)$$

where h is again the grid size. Thus, we have a second order approximation to the squared gradient.

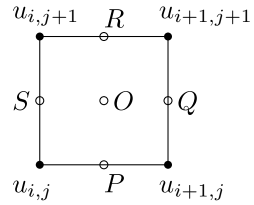

Figure 3: Illustration to diffusivity estimation from discrete samples. The u represent available data samples, the point O is the location where we need to evaluate the gradient magnitude.

Discretisation 2: Averaging after squaring. As an alternative to the preceding discretisation, one may also average the partial derivative approximations *after* squaring them:

$$\big(\partial_x u(O)\big)^2 \approx \frac{\big(\partial_x u(P)\big)^2 + \big(\partial_x u(R)\big)^2}{2}, \tag{50}$$

$$\big(\partial_y u(O)\big)^2 \approx \frac{\big(\partial_y u(S)\big)^2 + \big(\partial_y u(Q)\big)^2}{2}. \tag{51}$$

Each term in (50), (51) can be written using wavelet coefficients (see again Fig. 1):

$$\big(\partial_x u(P)\big)^2 = \big(-w_x^\alpha - w_{xy}^\alpha\big)^2, \qquad \big(\partial_y u(Q)\big)^2 = \big(-w_y^\alpha + w_{xy}^\alpha\big)^2, \tag{52}$$

$$\big(\partial_x u(R)\big)^2 = \big(-w_x^\alpha + w_{xy}^\alpha\big)^2, \qquad \big(\partial_y u(S)\big)^2 = \big(-w_y^\alpha - w_{xy}^\alpha\big)^2. \tag{53}$$

Putting (50)–(53) together, we obtain an alternative gradient estimation in the form

$$|\nabla u|^2(O) = \big(\partial_x u(O)\big)^2 + \big(\partial_y u(O)\big)^2 \approx \big(w_x^\alpha\big)^2 + \big(w_y^\alpha\big)^2 + 2\big(w_{xy}^\alpha\big)^2 \tag{54}$$

where all the three wavelet coefficient channels are included. Taylor expansions show that this discretisation comes down to

$$|\nabla u|^2 + \frac{1}{12}h^2(24u_{xy}^2 + u_x u_{xxx} + 3u_x u_{xyy} + 3u_y u_{xxy} + u_y u_{yyy}) + \mathcal{O}(h^4) \tag{55}$$

in point O. Although this expression has an additional error term in u_{xy}^2 compared to (49), it is an equally valid second order approximation to the squared gradient. If the sign of the other error terms is negative, the total discretisation error is even smaller than in the first discretisation.

Discretisation 3: Combining discretisations 1 and 2. We have seen that both equations (48) and (54) represent numerically consistent second order approximations to the squared gradient. Hence the same order of consistency holds true for any linear combination with weights summing up to 1. Thus, we may employ

$$|\nabla u|^2(O) = \big(\partial_x u(O)\big)^2 + \big(\partial_y u(O)\big)^2 \approx \big(w_x^\alpha\big)^2 + \big(w_y^\alpha\big)^2 + c \cdot \big(w_{xy}^\alpha\big)^2 \tag{56}$$

with any c. By choosing $c \geq 0$ one ensures that the approximation to $|\nabla u|^2$ can never become negative.

Practically, it seems that some care should be taken when combining the constant c in (56) and the weighting term q for the horizontal/vertical diffusion connectivity and the corresponding mixed-term shrinkage (44): If a coefficient channel w_{xy} is to be shrunken, it should get a chance to influence its fate by contributing to (56). However, the actually best weighting is signal-dependent, since the $\mathcal{O}(h^2)$ term in the discretisation error

$$\frac{1}{12}h^2(12cu_{xy}^2 + u_x u_{xxx} + 3u_x u_{xyy} + 3u_y u_{xxy} + u_y u_{yyy}) + \mathcal{O}(h^4) \qquad (57)$$

cannot be made rotationally invariant by choosing an appropriate value for c.

In the experiments below we use the value $c = 2$ in combination with the shrinkage rules (45)–(47). This comes down to Discretisation 2.

5 Rotation–Invariant Wavelet Shrinkage

In image analysis it is desirable that the features detected in the data do not depend on their orientation. A filter is called *rotation invariant* if its result is not influenced by a rotation of the input compensated by inverse rotation of the output. Unfortunately, classical wavelet shrinkage [13] is not invariant to rotation of the input data.

Several attempts to create wavelet transforms with improved rotation invariance have appeared in the literature, including the directional cycle spinning of Yu *et al.* [45], complex wavelets of Kingsbury [23], or the elaborated edgelet and curvelet transforms [37]. Some of these ideas are relatively difficult to implement or computationally significantly more complex than the traditional 2-D shift-invariant wavelet shrinkage from Section 2.3.

We have derived above the connection between 2-D discrete schemes for nonlinear diffusion on one hand, and shift-invariant 2-D Haar wavelet shrinkage on the other. The resulting coupled wavelet shrinkage rules (42)–(44) inherit a fundamental property from their diffusion origin: the rotation invariance of the nonlinear diffusion filter. It holds exactly for the grid size tending to zero, but we shall see that this property is also well approximated in realistic discrete situations with non-vanishing grid size. Our diffusion-inspired idea to improve the invariance to rotation by coupling the shrinkage of wavelet channels represents a very simple solution which hardly increases the computational complexity of a wavelet filter.

The diffusion–wavelet connection has been shown for a single iteration of a single-scale filter. In general, nonlinear diffusion is a single-scale iterative process, while wavelet shrinkage finds the solution using a single step on multiple scales. A hybrid multi-scale iterated filter seems to be a powerful and efficient alternative [15, 32]. It can be understood either as a nonlinear diffusion on the Laplacian pyramid of the signal [39], or as iterated shift-invariant wavelet shrinkage [5]. We apply the idea of coupled shrinkage to this general, iterated multiscale filter.

In the following experiments we compare the rotation invariance of two wavelet-based filters: the classical iterated shift-invariant 2-D wavelet shrinkage (18) with separate shrinkage of the coefficient channels, and the novel filter with shrinkage rules coupled according to (45)–(47). In all cases, we employ the

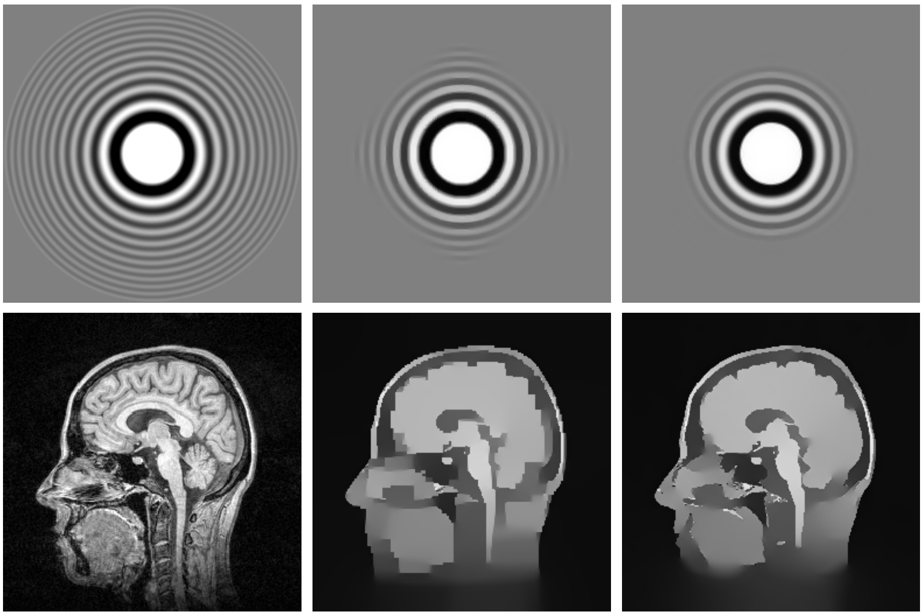

Figure 4: Experiments on rotation invariance. **Left:** input image. **Middle:** filtered with classical iterated shift invariant wavelet shrinkage. **Right:** method with channels coupled using (30)–(32). **Top:** ring image, five iterations on 8 levels of the wavelet decomposition. **Bottom:** head image, 100 iterations on 4 levels.

Haar wavelet basis combined with hard shrinkage. The wavelet decomposition is calculated on multiple scales.

In the first experiment, we start from the rotationally symmetric ring image (Fig. 4 top left). Examples of images obtained after 5 iterations of each method are seen in Fig. 4 top. One can observe that using the coupled shrinkage (45)–(47), the filtered result reveals a much better rotational symmetry. The difference between the two methods is further visualised on a medical image at the bottom of Fig. 4. At a comparable level of image simplification, the new method is able to avoid the blocky artifacts of the classical transform.

The graphs in Fig. 5 present a numerical evaluation of the errors in rotational symmetry of the filtered ring image. The measure of asymmetry was calculated as a sum of signal variances along circles of varied diameter, centered at the center of rotation of the input image. In agreement with the design principles, the rotation symmetry of both the single-step and the iterated filter with coupled channels is very good and by far outperforms the classical transform. Using channel coupling, the method can to a great extent overcome the limitations of the Haar basis.

6 Filtering of Colour Images

Greyscale images can be understood as functions $f\colon \Omega \to \mathbb{R}$. In some other cases, several physical quantities are measured at the same location in space: colour (RGB) images fall into this category, together with e.g. multi-echo mag-

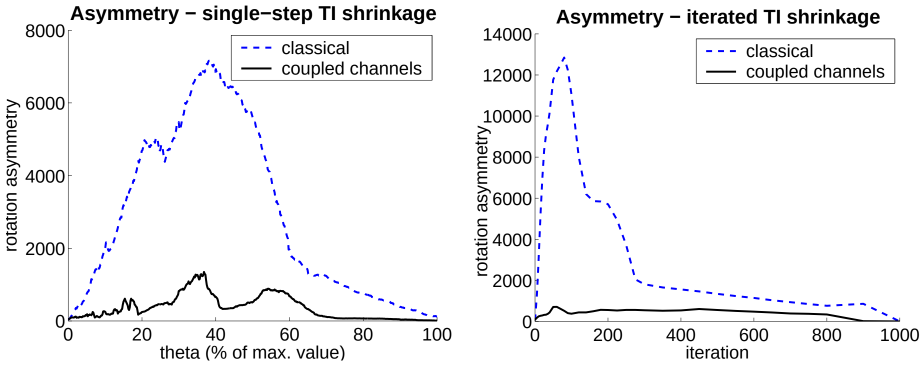

Figure 5: Evaluation of the errors in rotation symmetry of the filtered ring image. Single-step **(left)**, and iterated **(right)** shift-invariant wavelet shrinkage. On the left, 'theta' denotes the shrinkage parameter.

netic resonance medical data, multispectral LANDSAT measurements, and many others. In this case the function f maps Ω to $\mathbb{R}^m$, and the vector-valued data $f = (f_1, f_2, \ldots, f_m)^\top$ are represented by a collection of m scalar images.

For vector-valued images, the diffusion equation (19) translates into a set of equations

$$\partial_t u_1 = \operatorname{div}\left(g\Big(\sum_{j=1}^{m} |\nabla u_j|^2\Big)\, \nabla u_1\right) \tag{58}$$

$$\vdots$$

$$\partial_t u_m = \operatorname{div}\left(g\Big(\sum_{j=1}^{m} |\nabla u_j|^2\Big)\, \nabla u_m\right) \tag{59}$$

where the solution u is also composed of m images, $u = (u_1, u_2, \ldots, u_m)^\top$, with the initial condition $u_i = f_i$. Note that the diffusivity g depends on the entire vector u and is identical for all equations in the set (58)–(59). This has been adopted as a common practice for vector-valued diffusion (see e.g. [44, 42, 24]): In order to avoid inconsistencies between separate channels u_i, the equations (58)–(59) are coupled and the diffusion of individual images is synchronised through a common set of diffusivities g.

As described in [37], the typical approach to wavelet shrinkage of colour images is to convert the image first from RGB to YUV colour space, create the wavelet representation for each of the YUV components independently, shrink them independently and then compose the results. As we shall see in the experiments, this approach is better than independent filtering of RGB channels in that it reduces most of unpleasant colour misalignment. However, this separate YUV-based procedure tends to flatten the colour depth and variety of the filtered image, leading to a rather greyish look.

Having connected the diffusivity and shrinkage functions by (45)–(47), it is straightforward to extend the diffusion idea of coupling of colour channels to the wavelet domain. We first construct the wavelet representation for the R, G and B colour channels separately, which results in nine channels of wavelet coefficients, w_z^Z, $z \in \{x, y, xy\}$, $Z \in \{R, G, B\}$. Then the joint gradient of a

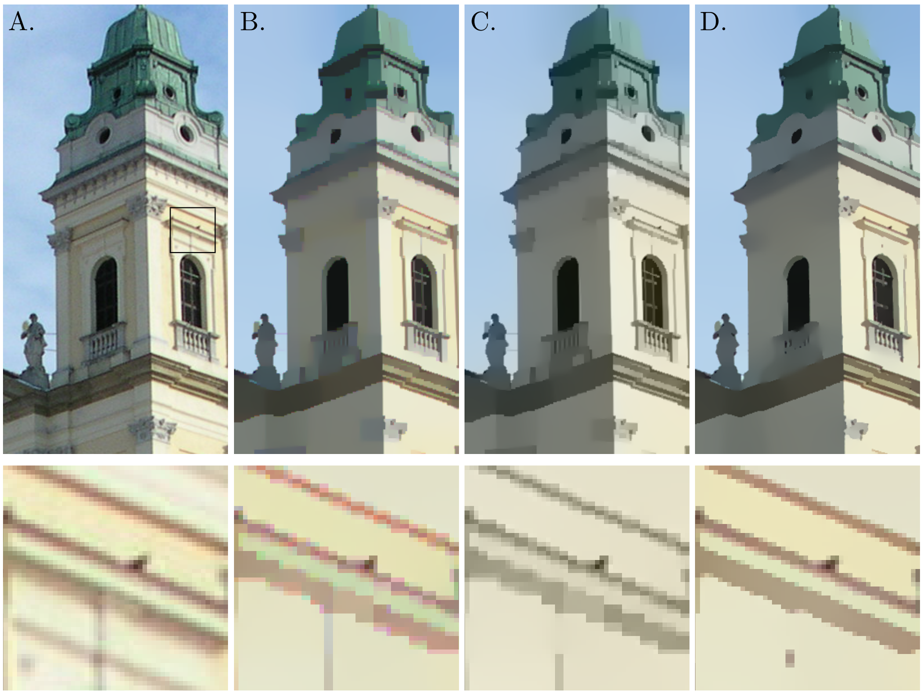

Figure 6: Wavelet shrinkage of colour images. Top: **A.** Input image.
B. Separate shrinkage of wavelet coefficients and separate RGB colour layers.
C. Separate shrinkage of wavelet coefficients and separate YUV colour layers.
D. Shrinkage with coupled coefficient channels and coupled RGB colour layers.
Bottom: details extracted from each image at the position marked in image A.

RGB colour image can be estimated using

$$\begin{aligned} |\nabla u|^2 &= |\nabla u_R|^2 + |\nabla u_G|^2 + |\nabla u_B|^2 \\ &= \left(w_x^R\right)^2 + \left(w_y^R\right)^2 + c \cdot \left(w_{xy}^R\right)^2 \\ &+ \left(w_x^G\right)^2 + \left(w_y^G\right)^2 + c \cdot \left(w_{xy}^G\right)^2 \\ &+ \left(w_x^B\right)^2 + \left(w_y^B\right)^2 + c \cdot \left(w_{xy}^B\right)^2 . \end{aligned} \tag{60}$$

Using this squared gradient magnitude to calculate the common diffusivity $\widetilde{g}$ for all channels, $\widetilde{g} = g(|\nabla u|^2)$, we are able to construct coupled wavelet shrinkage rules for colour images with both good rotation invariance and synchronised evolution of all colour channels:

$$S_\theta(w_z^Z) = w_z^Z\,(1 - 4\,\tau\,\widetilde{g}), \quad z \in \{x, y, xy\}, \quad Z \in \{R, G, B\}. \tag{61}$$

This procedure is applied at all positions and all scales of the wavelet representation.

Figures 6 and 7 show examples of iterated nondecimated Haar wavelet shrinkage of colour images. In both cases, some input image (A) is filtered using separate hard shrinkage of RGB colour channels (B), separate hard shrinkage of YUV colour channels (C), and the coupled hard shrinkage of RGB channels

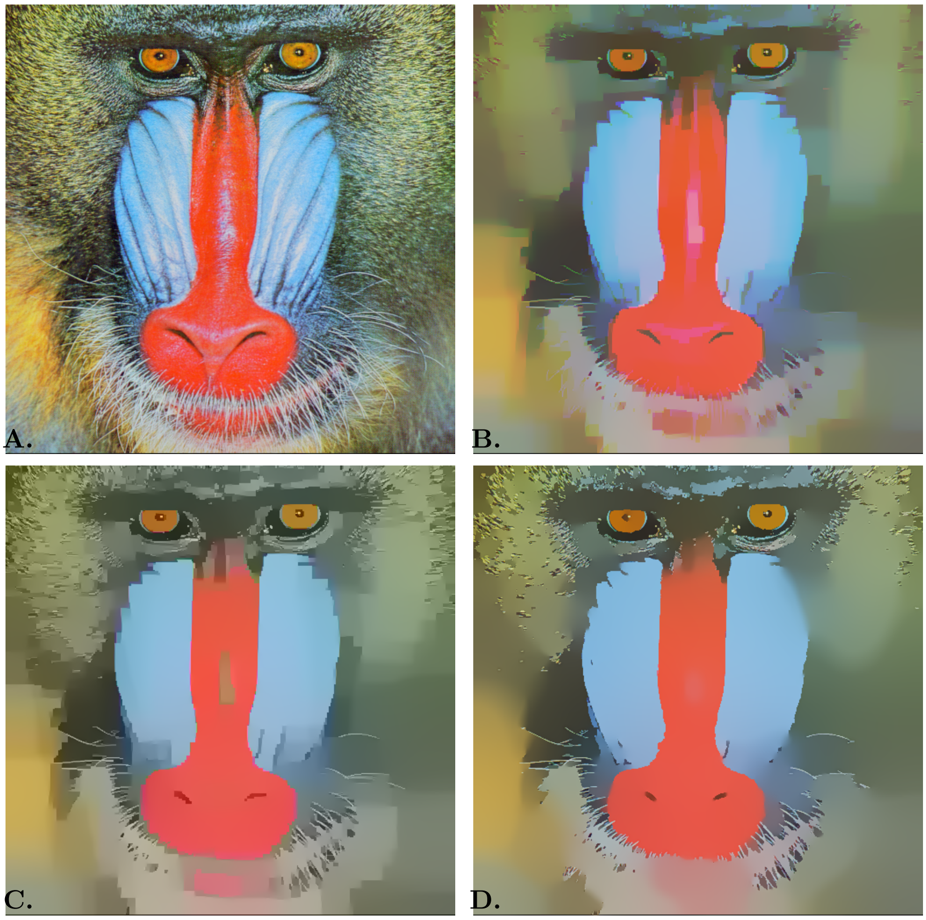

Figure 7: Wavelet shrinkage of colour images. **A.** Input image.
B. Separate shrinkage of wavelet coefficients and separate RGB colour layers.
C. Separate shrinkage of wavelet coefficients and separate YUV colour layers.
D. Shrinkage with coupled coefficient channels and coupled RGB colour layers.

according to (61) (D). The independent shrinkage of colour channels easily creates colour artifacts (B) or makes the image look greyish (C); both of them also contain some blocky artifacts resulting from the Haar basis employed. On the other hand, the coupled shrinkage (D) provides a more natural colour simplification, and at the same time exhibits good rotation invariance thanks to the synchronised shrinkage of the Haar wavelet coefficients. Our diffusion-inspired joint shrinkage also confirms recent results in [38] where a coupling of colour channels in the *Luv* space is studied for image enhancement.

7 Filtering of Tensor-Valued Data

Recently, matrix-valued data sets (so-called tensor fields) have gained importance in the field of image processing. This has been triggered by the following developments:

- Novel medical imaging techniques such as *diffusion tensor magnetic resonance imaging (DT-MRI)* have been introduced. DT-MRI is a 3-D imaging method that yields a diffusion tensor in each voxel. The resulting matrix field provides valuable information for brain connectivity studies as well as for multiple sclerosis or stroke diagnosis [34].

- Tensors have shown their use as a common description tool in image analysis, segmentation and grouping [20]. This also includes widespread applications of the so-called structure tensor (Förstner interest operator, second moment matrix, scatter matrix) [16] in fields ranging from motion analysis to texture segmentation.

- A number of scientific applications require to process tensor fields: The tensor concept is a common physical description of anisotropic behaviour in solid mechanics and civil engineering, where stress-strain relationships, inertia tensors, diffusion tensors, and permitivity tensors are used.

Since matrix-valued data are often polluted by noise, it would be desirable to develop wavelet shrinkage methods that remove noise without sacrifying discontinuities of the tensor fields. Moreover, it is clear that the image denoising method should be rotation invariant as well.

While several methods have been developed for denoising matrix-valued data, we are not aware of any wavelet-related concepts for matrix fields apart from Aldroubi's work on sampling in shift-invariant amalgam spaces [1]. Let us now describe how the connection between diffusion filtering and wavelet shrinkage can also be exploited in this situation.

For simplicity we consider some 2-D matrix field $F(x, y) = (f_{ij}(x, y))$ with $1 \leq i, j \leq 2$, but our reasoning also carries over to higher dimensions. In [41], a nonlinear diffusion filter for matrix-valued data is proposed that is based on the evolution

$$\partial_t u_{i,j} \;=\; \operatorname{div}\left(g\Big(\sum_{k,l} \nabla u_{k,l}^\top \nabla u_{k,l}\Big)\, \nabla u_{i,j}\right) \qquad \forall i, j. \tag{62}$$

It is evident that methods of this type are rotationally invariant in their continuous formulation.

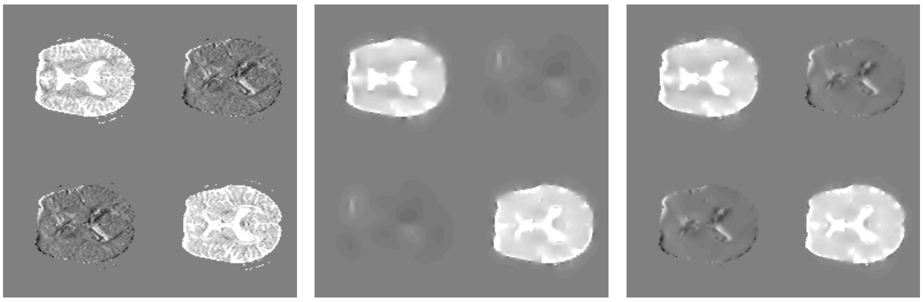

Figure 8: Filtering of tensor-valued images. **Left:** input data, symmetric 2×2 tensor at each position, visualised by tiling the components into a single image. **Middle:** after filtering with classical shift-invariant Haar wavelet shrinkage. **Right:** Haar shrinkage, shift- and rotation-invariant with coupled evolution of the tensor components. Both results used 4 spatial levels of the wavelet decomposition, 1 step, and a hard threshold of 50.

All we have to do now is to transfer these diffusion results to the corresponding wavelet setting. In the same way as for filtering of vector-valued images, we can use shrinkage functions of type

$$S_\theta(w_z^{i,j}) = w_z^{i,j}\,(1 - 4\,\tau\,\widetilde{g}), \qquad z \in \{x, y, xy\} \tag{63}$$

where the joint diffusivity $\tilde{g}$ satisfies

$$\widetilde{g} := g\Big(\sum_{i,j}\Big(\big(w_x^{i,j}\big)^2 + \big(w_y^{i,j}\big)^2 + c\cdot\big(w_{xy}^{i,j}\big)^2\Big)\Big). \tag{64}$$

In Figure 8 this joint shrinkage is compared with a shift-invariant shrinkage where all wavelet coefficients are shrunken independently and no channel coupling takes place. This figure depicts a 2-D section from a three-dimensional DT-MRI data set. We observe that the coupled shrinkage offers a significantly better preservation of discontinuities in the low-contrast off-diagonal channels.

8 Conclusions

In this paper we have investigated connections between fully discrete nonlinear diffusion filters and shift-invariant Haar wavelet shrinkage in which the idea of *coupling* plays a central role. While the coupling concept is very common in diffusion filtering, hardly any attention has been paid to it in the context of wavelet shrinkage. By transferring coupling ideas to this area, we have been able

- to introduce wavelet shrinkage methods with significantly improved rotation invariance. They benefit from the rotation invariance of continuous nonlinear diffusion filtering.
- to investigate techniques for colour wavelet shrinkage in which the shinkage of the channels is synchronized. This avoids the creation of colour artifacts and feature misalignments that are common for separate colour channel filtering in RGB or YUV space.

- to propose a wavelet shrinkage technique for matrix-valued images in which the channel evolution is synchronised. Also in this case the method performs better than separate channel shrinkage.

It should be noted that all this coupled wavelet shrinkage strategies are hardly more complicated than separate shrinkage, while offering significantly better performance. We have shown that even for the simple Haar basis, very satisfactory results are possible.

In order to keep things as simple as possible, we have restricted ourselves to the 2-D case in the present paper. However, it is evident that the proposed concept is generic and can be generalised in a straightforward way to any arbitrary dimension.

In our ongoing work we are investigating connections between PDE-based filters and wavelet shrinkage also in the case of more advanced wavelets with a higher number of vanishing moments. First results in this direction will be reported in [43].

Acknowledgements

Our research was funded by the project *Relations between Nonlinear Filters in Digital Image Processing* within the DFG Priority Programme 1114 *(Mathematical Methods for Time Series Analysis and Digital Image Processing)*. This is gratefully acknowledged. We also thank Ron Kimmel (Technion, Haifa) and Nir Sochen (Tel Aviv University) for discussions on colour wavelet shrinkage, and Anna Vilanova i Bartrolí (TU Eindhoven) and Carola van Pul (Maxima Medical Center, Eindhoven) for providing us with the DT-MRI data set. The experiments were completed with the contribution of Varadarajulu Reddy Pyda and Srikanth Ramaka.

References

[1] A. Aldroubi and P. Basser. Reconstruction of vector and tensor fields from sampled discrete data. In D. Larson and L. Baggett, editors, *The Functional and Harmonic Analysis of Wavelets*, volume 247 of *Contemporary Mathematics*, pages 1–15. AMS, Providence, 1999.

[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] 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.

[4] 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, Mar. 1998.

[5] 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.

[6] 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, Sept. 2000.

[7] A. Cohen, W. Dahmen, I. Daubechies, and R.DeVore. Harmonic analysis in the space BV. Technical Report 195, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Germany, Sept. 2000.

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

[9] 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.

[10] 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.

[11] 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.

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

[13] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrica*, 81(3):425–455, 1994.

[14] S. Durand and J. Froment. Reconstruction of wavelet coefficients using total-variation minimization. *SIAM Journal on Scientific Computing*, 24(5):1754–1767, 2003.

[15] A. K. Fletcher, K. Ramchandran, and V. K. Goyal. Wavelet denoising by recursive cycle spinning. In *Proc. 2002 IEEE International Conference on Image Processing*, volume 2, pages II: 873–876, 2002.

[16] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987.

[17] A. S. Frangakis, A. Stoschek, and R. Hegerl. Wavelet transform filtering and nonlinear anisotropic diffusion assessed for signal reconstruction performance on multidimensional biomedical data. *IEEE Transactions on Biomedical Engineering*, 48(2):213–222, 2001.

[18] H. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7(4):469–488, 1998.

[19] H. Gao and A. G. Bruce. WaveShrink with firm shrinkage. *Statistica Sinica*, 7:855–874, 1997.

[20] G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision*. Kluwer, Dordrecht, 1995.

[21] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian. A real-time algorithm for signal analysis with the help of the wavelet transform. In J. M. Combes, A. Grossman, and P. Tchamitchian, editors, *Wavelets: Time-Frequency Methods and Phase Space*, pages 286–297. Springer, Berlin, 1989.

[22] S. L. Keeling and R. Stollberger. Nonlinear anisotropic diffusion filters for wide range edge sharpening. *Inverse Problems*, 18:175–190, Jan. 2002.

[23] N. G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Journal of Applied and Computational Harmonic Analysis*, 10(3):234–253, 2001.

[24] Z. Krivá and K. Mikula. An adaptive finite volume method in processing of color images. In *Proc. Algoritmy 2000 Conference on Scientific Computing*, pages 174–187, Podbanské, Slovakia, Sept. 2000.

[25] 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.

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

[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, Berlin, 2003. Springer.

[30] P. Mrázek, J. Weickert, and G. Steidl. Correspondences between wavelet shrinkage and nonlinear diffusion. In L. D. Griffin and M. Lillholm, editors, *Scale-Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 101–116, Berlin, 2003. Springer.

[31] P. Mrázek, J. Weickert, and G. Steidl. Diffusion-inspired shrinkage functions and stability results for wavelet denoising. Technical Report 96, Dept. of Mathematics, Saarland University, Saarbrücken, Germany, Oct. 2003. To appear in *International Journal of Computer Vision*.

[32] P. Mrázek, J. Weickert, G. Steidl, and M. Welk. On iterations and scales of nonlinear filters. In O. Drbohlav, editor, *Proc. Eighth Computer Vision Winter Workshop*, pages 61–66, Valtice, Czech Republic, Feb. 2003. Czech Pattern Recognition Society.

[33] 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.

[34] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3):637–648, Dec. 1996.

[35] J. Shen. A note on wavelets and diffusion. *Journal of Computational Analysis and Applications*, 5(1):147–159, 2003.

[36] J. Shen and G. Strang. On wavelet fundamental solutions to the heat equation – heatlets. *Journal of Differential Equations*, 161(2):403–421, 2000.

[37] J. Starck, E. Candès, and D. L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6):670–684, 2002.

[38] J. Starck, F. Murtagh, E. Candès, and D. L. Donoho. Gray and color image contrast enhancement by the curvelet transform. *IEEE Transactions on Image Processing*, 12(6):706–717, June 2003.

[39] 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(2):686–713, 2004.

[40] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley–Cambridge Press, Wellesley, 1996.

[41] D. Tschumperlé and R. Deriche. Diffusion tensor regularization with constraints preservation. In *Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 948–953, Kauai, HI, Dec. 2001. IEEE Computer Society Press.

[42] J. Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3–4):199–210, Mar. 1999.

[43] J. Weickert, G. Steidl, P. Mrázek, M. Welk, and T. Brox. Diffusion filters and wavelets: What can they learn from each other? In N. Paragios, Y. Chen, and O. Faugeras, editors, *The Handbook of Mathematical Models in Computer Vision.* Springer, New York, 2005. In press.

[44] R. Whitaker and G. Gerig. Vector-valued diffusion. In B. M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, volume 1 of *Computational Imaging and Vision*, pages 93–134. Kluwer, Dordrecht, 1994.

[45] T. Yu, A. Stoschek, and D. Donoho. Translation- and direction-invariant denoising of 2-D and 3-D images: experience and algorithms. In M. A. Unser, A. Aldroubi, and A. F. Laine, editors, *Wavelet Applications in Signal and Image Processing IV*, volume 2825 of *Proceedings of SPIE*, pages 608–619. SPIE Press, Bellingham, 1996.