

Nonlinear Regularized Reaction-Diffusion Filters for Denoising of Images with Textures

GERLIND PLONKA¹ AND JIANWEI MA²

¹ Department of Mathematics, University of Duisburg-Essen, Campus Duisburg,
47048 Duisburg, Germany
gerlind.plonka@uni-due.de

² School of Aerospace, Tsinghua University, Beijing 100084, China
jma@tsinghua.edu.cn

Abstract

Denoising is always a challenging problem in natural imaging and geophysical data processing. In this paper we consider the denoising of texture images using a nonlinear reaction-diffusion equation and directional wavelet frames. In our model, a curvelet shrinkage is used for regularization of the diffusion process to preserve important features in the diffusion smoothing and a wave atom shrinkage is used as the reaction in order to preserve and enhance interesting oriented textures. We derive a digital reaction-diffusion filter that lives on graphs and show convergence of the corresponding iteration process. Experimental results and comparisons show very good performance of the proposed model for texture-preserving denoising.

Key words. reaction-diffusion, second-generation curvelets, wave atoms, digital TV, regularization, denoising

1 Introduction

Denoising is one important operation in image processing. Most denoising methods assume that the image is piecewise smooth while the noise is a high-frequency oscillation. So one tries to remove the oscillations using local or adaptive smoothing strategies. Such methods involve anisotropic diffusion, see e.g. [12, 39, 45, 50], means algorithms, see e.g. [7, 41], regularization techniques, see e.g. [15, 19, 24, 33, 42], or wavelet systems, see e.g. [17, 28, 43], etc. But actually, this assumption is often not suitable for natural images, e.g., many fine structures such as textures are as oscillatory as noise. The texture will unfortunately be treated as being noisy and will be wiped away. For natural images, one is almost unable to obtain satisfying results only using one method due to the complex structures and irregular details of images.

In [15], Chan, Osher and Shen introduced a nonlinear digital TV filter that can be interpreted as a translation of the classical analog TV restoration model invented by Rudin, Osher and Fatemi [42] to the digital case. The digital TV filter is a data-dependent filter, capable of denoising data without blurring jumps and edges for non-flat images. Different from most statistical filters, it has a simple fixed filter structure and an exact formula for the filter coefficients that intrinsically encode the edge information. It is simple to understand by readers who do not have knowledge on PDE's and numerical approximations.

In this paper, we propose a very general regularized reaction-diffusion model, that aims to improve the reconstruction of images with oriented textures. In the spirit of the digital TV-filter in [15], we introduce a discrete filter living on graphs associated to our continuous model. Therefore the proposed regularized digital reaction-diffusion filter is also useful for non-flat features.

We also present the analysis of the new digital filter and prove convergence for suitable regularization parameters $\lambda \geq 0$ balancing the diffusion and the reaction term.

In Section 2 we introduce the continuous reaction-diffusion model that applies a curvelet shrinkage as a regularization of the diffusion process and a wave atom shrinkage in the reaction part. We shall give detailed arguments for the special choice of that model and a comprehensive survey of related work. The special choice of the operators in the regularized diffusion and in the reaction term is the key for a successful separation of texture and noise in images. In Section 3 we introduce the corresponding new discrete reaction-diffusion filter on graphs. Section 4 is devoted to the study of the properties and the convergence of the proposed discrete reaction-diffusion filter. Further, we show that the limit of the iteration process indeed satisfies a digital analogue of certain Euler equations. Hence the filter can also be derived from a regularization method. In Section 5 we show the performance of the new filter in comparison with other denoising methods. Finally, we draw a conclusion in Section 6.

We note that this paper is part II of the combination of curvelets and nonlinear diffusion. In our part I [32], a projected TV diffusion is applied as a postprocessing to suppress the pseudo-Gibbs and curvelet-like artifacts.

2 Reaction-diffusion model

Let $u_0(x)$ be the noise contaminated version of an image $u(x)$, i.e.,

$$u_0(x) = u(x) + n(x).$$

Here $n(x)$ denotes random noise with mean 0 and variance σ^2 ,

$$En(x) = 0, \quad En^2(x) = \sigma^2.$$

Further, we suppose that the original image contains also texture parts which need to be preserved during the denoising process.

We now propose a reaction-diffusion model of the form

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla u|) \nabla u) + \gamma(Su_0 - u) \quad (2.1)$$

or more generally

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla(P_\sigma u)|) \nabla u) + \gamma(Su_0 - u) \quad (2.2)$$

with the original noisy signal u_0 as initial condition, and with homogeneous Neumann boundary conditions. Here $|\nabla u(x, y)| := \sqrt{u_x(x, y)^2 + u_y(x, y)^2}$ denotes the Euclidean norm of the partial derivatives of u . The first term on the right-hand side is a diffusion term as described in [12], using a certain regularization $P_\sigma u$. Typically, $g(|x|)$ is a non-negative decreasing function with $\lim_{x \rightarrow \infty} g(|x|) = 0$. The diffusivity g controls the smoothing process by admitting strong diffusion if the gradient ∇u is small (possibly caused by

noise) and by slowing down (or even stop) the smoothing for large gradients. In this paper, we restrict our attention to bounded diffusivities g . Without loss of generality let

$$0 < g(|x|) \leq 1, \quad x \in \mathbb{R}.$$

This restriction is e.g. satisfied for the Perona-Malik diffusivity $g(|x|) := (1 + x^2/\kappa^2)^{-1}$, the Charbonnier diffusivity $g(|x|) := (1 + x^2/\kappa^2)^{-1/2}$, and the truncated TV diffusivity $g(|x|) := \min\{1, \frac{1}{|x|}\}$, [39, 48].

As regularization operator P_σ we suggest to take here the curvelet-shrinkage

$$P_\sigma u = T^{-1}\theta T(u),$$

where T denotes the curvelet transform and T^{-1} the inverse transform [8, 9, 10]. For a short summary on the curvelet transform we refer to [32]. Further, θ can be taken as a soft threshold function defined by a fixed threshold $\sigma > 0$,

$$\theta_s(x) = \begin{cases} x - \sigma, & x \geq \sigma, \\ 0, & |x| < \sigma, \\ x + \sigma & x < -\sigma, \end{cases}$$

or a hard threshold function

$$\theta_h(x) = \begin{cases} x, & |x| \geq \sigma, \\ 0, & |x| < \sigma. \end{cases}$$

As we will see in Section 4, the convergence of the proposed discretization of (2.2) is based on the assumption that P_σ is a continuous operator. Therefore, instead of θ_h the continuous garrote threshold

$$\theta_g(x) := \begin{cases} x - \frac{\sigma^2}{x}, & |x| \geq \sigma, \\ 0, & |x| < \sigma. \end{cases}$$

may be a good choice, where large coefficients nearly remain unaltered.

The second term on the right-hand side of (2.1) resp. (2.2) is the reaction term for enhancement of oriented textures, consisting of the difference of the processed noisy image Su_0 and u . Here the nonlinear operator S should preserve and enhance the important features of the image. One may take here again a wavelet shrinkage or a curvelet shrinkage, i.e.,

$$Su_0 = T^{-1}\theta_h T(u_0).$$

We especially propose to use a transform based on wave atoms (see [21]). In fact, the reaction-diffusion model (2.2) with wave atoms has been applied in [31]. In this case, the reaction term is of the form

$$(WA)^{-1}\theta(WA)(u_0) - u$$

where WA denotes the transform of wave atoms (see [21]).

2.1 Why use the curvelets as a regularization of diffusion?

The well-known Perona-Malik model (obtained for $\gamma = 0$ in (2.1)) has many desirable properties, and different numerically stable schemes have been proposed for image denoising based on this model [48, 49, 50]. However from analytical point of view the Perona-Malik model [39] is a notoriously ill-posed problem [22, 23, 26]. Another drawback of the PM diffusion is its sensitivity to noise. Noise often introduces very large oscillations of the

gradient ∇u , therefore the gradient-based model possibly misjudges true edges and heavy noise leading to undesirable diffusion in regions where there is no true edge. In [12], a Gaussian regularization is proposed, i.e.,

$$P_\sigma u(x) = (K_\sigma * u)(x) = \int_{\mathbb{R}^2} K_\sigma(x-y) u(y) dy, \quad K_\sigma := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x|^2}{2\sigma^2}\right).$$

Then, considering the model (2.2) with $\gamma = 0$, existence, uniqueness and regularity of a solution has been established for $\sigma > 0$ in [12, 45]. However, Gaussian filtering blurs edges and finer textures. This behavior seems somewhat against the purpose of the PM equation (i.e. sharpen the edges). Recently, a wavelet regularization has been also considered [51].

But the conventional tensor-product 2D wavelets have not good performance at representing line singularities because they ignore the geometric properties of objects and do not exploit the regularity of edge curves.

The curvelet transform [8, 9, 10] is a new geometric multiscale transform, which allows an optimally sparse representation of objects with C^2 -singularities. The needle-shape elements of this transform own very high directional sensitivity and anisotropy (see Fig. 1 (left)). For a smooth object f with discontinuities along smooth curves, the best m -term approximation \tilde{f}_m by curvelet threshold obeys $\|f - \tilde{f}_m\|_2^2 \leq Cm^{-2}(\log m)^3$, while for wavelets the decay rate is only m^{-1} . It has been shown that the new transform represents edges and singularities along curves much more efficiently than traditional wavelet transforms, e.g. [29, 30, 43].

The motivation of our regularized model is to improve the ill-posed problem of diffusion by using the second-generation curvelet transform instead of traditional Gaussian filtering. The curvelet pre-processing can effectively remove the noise while preserving the edges well, leading to better discontinuity-preserving diffusion. Since the curvelet shrinkage works like a smoothing kernel applied to u , the ideas of Catté, Lions, Morel and Coll [12] can be also applied to our model, and the proof of existence and uniqueness is not altered by the additional reaction term. This fact has been stated already for the regularized Nordström model in [12], p. 185.

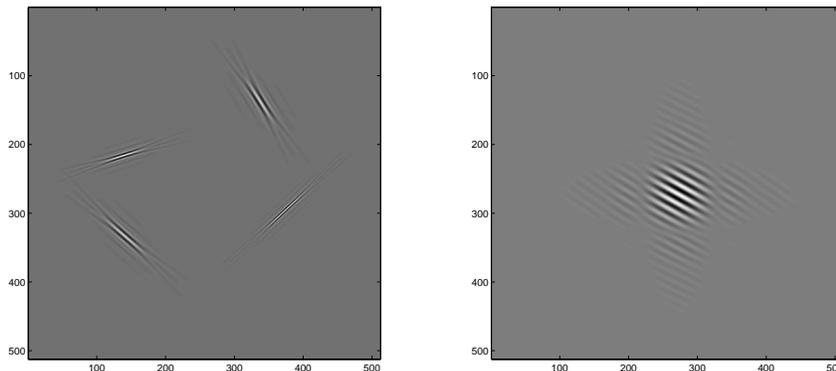


Figure 1: Elements of curvelets (left) and wave atoms (right) in spatial domain.

2.2 Why use wave atoms in the reaction term?

The motivation to use the reaction term $Su_0 - u$ is that the model can preserve even enhance the wanted oriented textures.

In [1], Acton et al. achieved an oriented texture completion by an AM-FM reaction-diffusion method using Gabor filters as reaction term. Very recently, Demanet and Ying [21] introduced so-called wave atoms, that can be seen as a variant of 2D wavelet packets and obey the parabolic scaling of curvelets $wavelength = (diameter)^2$ (see Fig. 1 (right)). Oscillatory functions or oriented textures (e.g., fingerprint, seismic profile, engineering surfaces) have a significantly sparser expansion in wave atoms than in other fixed standard representations like Gabor filters, wavelets, and curvelets.

Wave atoms have the ability to adapt to arbitrary local directions of a pattern, and to sparsely represent anisotropic patterns aligned with the axes. In comparison to curvelets, wave atoms not only capture the coherence of the pattern along the oscillations, but also the pattern across the oscillations.

It is therefore natural to apply the wave-atoms shrinkage as a reaction for enhancement of the oriented textures.

2.3 Related work

The idea of using reaction diffusion equations, combining the diffusion for noise filtration and the reaction to improve the contrast, can already be found in [1, 2, 5, 16, 18, 40].

The model in [18] is of the form

$$\frac{\partial u}{\partial t} = \sigma \epsilon^2 \nabla \cdot (A_\epsilon(u) \nabla u) + f(u) \quad \text{in } \Omega \subset \mathbb{R}^2$$

with initial condition $u(\cdot, 0) = u_0$ in Ω and $u = 0$ on $\delta\Omega$. The nonlinear operator $A_\epsilon(u)$ approximates the orthogonal projection of u onto the direction which is perpendicular to the gradient of u_ϵ . Here the function u_ϵ coincides with u for $x \in \Omega$ and $d(x, \delta\Omega) > 2\epsilon$, and $u_\epsilon = 0$ on $\delta\Omega$. The reaction term $f \in C^1$ satisfies $f(\pm 1) = 0$, $x f(x) > 0$ for $x \in (-1, 1) \setminus \{0\}$. One problem occurring in the model of [18] is that the parameter ϵ should be small such that u_ϵ well approximates u , but this leads to a small amount of diffusion.

The reaction-diffusion model in [2] is based on the equation

$$\frac{\partial u}{\partial t} = g(|G * \nabla u|) \|\nabla u\| \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right) + f(u)$$

with the initial condition $u(0, x) = u_0(x)$. The diffusivity function $g(x)$ is a nonnegative, nonincreasing function with $g(0) = 1$, G is a convolution kernel, and f is a Lipschitz function with a finite number of zeros. Compared with (2.2), the regularized diffusivity function works here as a factor (as usual for implicit snakes) and is not directly incorporated in the diffusion term. The term $\|\nabla u\| \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right)$ diffuses u in the orthogonal direction to the gradient ∇u while the term $g(|G * \nabla u|)$ controls the diffusion speed. The function f determines the asymptotic state of the equation based on a Lloyd quantizer. In [16], a similar model has been investigated, where $f(u)$ is replaced by $\nabla g(\nabla(G_\sigma * u)) \cdot \nabla u$.

In [40] the simplified diffusion-reaction equation

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla u|) \nabla u) + \gamma(u_0 - u)$$

with $u(0, x) = u_0(x)$ and homogeneous Neumann boundary conditions has been considered and discretized on graphs. This equation coincides with (2.1) for $S = Id$. In spirit of

the mechanism, various other reaction diffusion equations have been proposed for image processing without rigorous proofs of existence and uniqueness of solutions [1, 5].

The considered reaction-diffusion model (2.1) is also closely related to the minimization of energy functionals.

In [36], Nordström suggested to obtain a reconstruction u of a degraded image u_0 by minimizing the energy functional

$$E(u, \omega) := \int_{\Omega} \gamma(u_0 - u)^2 + \omega |\nabla u|^2 + \kappa^2(\omega - \ln \omega) \quad (2.3)$$

where the parameters γ and κ are positive weights and $\omega = \omega_u : \Omega \rightarrow [0, 1]$ gives a fuzzy edge representation. In regions with small variations of u the diffusivity function ω approaches 1 while at edges ω it is close to 0. The first term of E punishes deviations of u from u_0 , the second term detects unsmoothness of u and the last measures the extend of edges. The corresponding Euler equations to this energy functional are given by

$$\begin{aligned} 0 &= \gamma(u_0 - u) + \nabla \cdot (\omega \nabla u), \\ 0 &= \kappa^2 \left(1 - \frac{1}{\omega}\right) + |\nabla u|^2. \end{aligned}$$

with homogeneous Neumann boundary conditions. The second equation leads to the Perona-Malik diffusivity

$$\omega = \left(1 + \frac{|\nabla u|^2}{\kappa}\right)^{-1},$$

and we obtain

$$\nabla \cdot (g(|\nabla u|) \nabla u) + \gamma(u_0 - u) = 0$$

with $g(|\nabla u|) = \omega$. This is the steady-state equation of (2.1) for Perona-Malik diffusivity g and the operator $S = Id$. Observe that the energy functional E is nonconvex. Therefore, it can possess numerous local minima. Indeed, Nordström was not able to establish convergence of his discretized minimizing procedure for (2.3) for small γ .

In the original ROF model developed by Rudin, Osher and Fatemi [42], the wanted denoising is obtained by minimizing the energy

$$E(u) := |u|_{BV} + \lambda \|u_0 - u\|_{L^2(\Omega)}^2,$$

where λ is an fitting parameter and $|u|_{BV}$ denotes the BV seminorm in the space $BV(\Omega)$ of functions with bounded variation. For a rigorous definition of the BV-seminorm

$$|u|_{BV} = \int_{\Omega} |\nabla u| = \sup_{|g| \leq 1, g \in C_c^1(\Omega)} \int_{\Omega} u(\nabla \cdot g)$$

with $|g| = \sqrt{g_1^2 + g_2^2}$ we refer to [33].

The ROF model has been analyzed by several authors, see e.g. [13, 35]. One of its drawbacks is the so-called staircasing effect [35].

Many regularization methods have been devised by designing $\Phi(|\nabla u|)$ instead of $|\nabla u|$ in the ROF model, see e.g. [6, 13, 14, 34, 19]. These techniques essentially take the local smooth constraint or statistical behavior as a priori knowledge. Different improvements of the ROF model aim to achieve better behavior especially with respect to oscillating patterns by replacing the norm of the fidelity term $\|u - u_0\|_{L^2(\Omega)}$ by a more suitable norm [3, 20, 25, 27, 33, 37, 38, 46].

The Euler equations that are related to the minimization problems can be regarded as the steady state of suitable reaction-diffusion equations. In particular, for the ROF model we obtain the Euler equations

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u_0 - u) = 0,$$

where the expression $\nabla u/|\nabla u|$ needs to be defined suitably for $|\nabla u| = 0$. This can be easily done, see [33]. The steepest descent marching gives

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u_0 - u) \quad (2.4)$$

with $u(x, 0) = u_0$ and Neumann boundary conditions. In particular, replacing the term $(|\nabla u|)^{-1}$ by a more general diffusivity function $g(|\nabla u|)$, we obtain (2.1) with $S = Id$.

For better characterization of texture, Meyer [33] suggested (in the special case $\Omega = \mathbb{R}^2$) to replace the L^2 -norm of $u_0 - u$ in the functional $E(u)$ by $\|f - u\|_{BV^*}$, where $\|\cdot\|_{BV^*}$ denotes the norm of a space that can be seen as a dual space of $BV(\Omega)$. His model is based on the observation that a piecewise smooth image u belongs to the space $BV(\mathbb{R}^2)$ while a texture image v belongs to a different family of functions denoted by $BV^*(\mathbb{R}^2)$. This notation implies the existence of $g = (g_1, g_2)^T$ with $g_1, g_2 \in L^\infty(\mathbb{R}^2)$ such that $v(x, y) = \nabla \cdot g(x, y) = \partial_x g_1(x, y) + \partial_y g_2(x, y)$ and the BV^* norm is then defined by

$$\|v\|_{BV^*} = \|(|g_1|^2 + |g_2|^2)^{1/2}\|_\infty.$$

An approximation of the Meyer model due to Vese and Osher [46] is of the form

$$F(u, g) = \left(\int_\Omega \lambda(u_0 - u - \nabla \cdot g)^2 + |\nabla u| \right) + \mu \left(\int_\Omega |g|^p \right)^{1/p}$$

with $p \geq 1$ and $\lambda, \mu > 0$, thereby finding the separation $f = u + v$. Other ideas to replace the L^2 -norm of $u_0 - u$ can be found in [3, 4, 20, 27, 38].

In contrast with these approaches, we propose another way here by replacing $\|u_0 - u\|_{L^2}^2$ by $\|Su_0 - u\|_{L^2}^2$, where S already carries out a denoising based on harmonic analysis methods. The shrinkage results by wave atoms are taken as a pseudo-observation in the model. Another approach for texture separation that combines a sparse representation of the image by special function systems (wavelets, Gabor frames, etc) with a regularization method can be found in [44]. The core idea there is to choose two appropriate dictionaries, one for representation of texture, and the other for the natural scene parts.

3 Discrete model on graphs

We want to discretize the model (2.2) as follows. For time discretization, we apply a semi-implicit scheme with step size τ . Using $t_j := j\tau$, $j \in \mathbb{N}_0$ and $u^j(x)$ as an approximation of $u(x, t_j)$, we obtain

$$\frac{u^{j+1} - u^j}{\tau} = \nabla \cdot (g(|\nabla P_\sigma u^j|) \nabla u^j) + \gamma(Su^0 - u^{j+1})$$

leading with $\lambda := \tau\gamma$ to

$$u^{j+1} = \frac{1}{1 + \lambda} (u^j + \lambda Su^0 + \tau \nabla \cdot (g(|\nabla P_\sigma u^j|) \nabla u^j)). \quad (3.1)$$

Let us remark that the explicit time discretization of the model (2.2)

$$u^{j+1} = u^j + \lambda(Su^0 - u^j) + \tau \nabla \cdot (g(|\nabla P_\sigma u^j|) \nabla u^j)$$

does not lead to a stable scheme for large λ .

For spatial discretization consider noisy data living on graphs. See also [15, 24] and references therein for similar graph models. For that purpose we need to introduce some notations.

A digital domain is modeled by a graph $[\Omega, E]$ with a finite set $\Omega \subset \mathbb{R}^N$ of D nodes and an edge dictionary E . If $\alpha, \beta \in \Omega$ are linked by an edge e , we write $\alpha \sim \beta$ as well as $\alpha \prec e$ and $\beta \prec e$. Throughout the paper, we suppose that the graph $[\Omega, E]$ is connected, i.e., each node $\alpha \in \Omega$ is endpoint of at least one edge.

Let a digital signal u be a function on Ω , $u : \Omega \rightarrow \mathbb{R}$. We can assign a linear order to all nodes of Ω , $\alpha_1 < \alpha_2 < \dots < \alpha_D$. The value at node α is denoted by u_α . Then u is completely characterized by the vector $\mathbf{u} = (u_1, \dots, u_D)^T \in \mathbb{R}^D$. Let $e = e_{\alpha, \beta}$ denote the edge between α and β . Then the length of e is given by the Euclidean norm $\|e\| := \|\alpha - \beta\|_2$. We assume, the lengths of the edges in E are normalized in a way that $\min_{e \in E} \|e\| = 1$. The edge derivative of u along $e = e_{\alpha, \beta}$ at α is now given by

$$D_e u = \frac{\partial \mathbf{u}}{\partial e} \Big|_\alpha := \frac{u_\beta - u_\alpha}{\|e_{\alpha, \beta}\|}, \quad |D_e u| = |D_{\alpha, \beta}| := \frac{|u_\alpha - u_\beta|}{\|e_{\alpha, \beta}\|}.$$

Obviously, we have $\frac{\partial \mathbf{u}}{\partial e} \Big|_\alpha = -\frac{\partial \mathbf{u}}{\partial e} \Big|_\beta$.



Figure 2: Typical partial graphs for rectangular image domains.

We start now with a given noisy signal $\mathbf{u}^0 = (u_\alpha^0)_{\alpha \in \Omega}$ and use $\mathbf{v}^0 := \mathbf{S}\mathbf{u}^0$ as well as $\tilde{\mathbf{u}}^j := \mathbf{P}_\sigma \mathbf{u}^j$. Here, we assume that \mathbf{S} and \mathbf{P}_σ are suitable discrete approximations of the corresponding operators on $L^2(\Omega)$.

In particular, using an orthogonal or biorthogonal wavelet transform defining the operator $\mathbf{S} : \mathbb{R}^D \rightarrow \mathbb{R}^D$, it can be written in matrix form

$$\mathbf{S}\mathbf{u}^0 = \mathbf{v}^0 = \tilde{\mathbf{M}}^T \Theta_h(\mathbf{M}\mathbf{u}^0) \mathbf{M}\mathbf{u}^0,$$

where \mathbf{M} and $\tilde{\mathbf{M}}^T$ are the corresponding transform matrices (see e.g. [11]) and $\Theta_h(\mathbf{M}\mathbf{u}^0)$ is a diagonal matrix representing the hard threshold, i.e., the diagonal entries of $\Theta_h(\mathbf{M}\mathbf{u}^0)$ are 1 or 0, depending on the size of the corresponding component in the vector $\mathbf{M}\mathbf{u}^0$. Also, the discrete curvelet transform can be written in this form (see [10]), where the transform matrices \mathbf{M} and $\tilde{\mathbf{M}}$ need not to be quadratic, but are of the same size. In [10], a numerical procedure for the discrete curvelet transform is described based on the two-dimensional fast Fourier transform and using $\mathcal{O}(D(\log D)^2)$ arithmetical operations. An algorithm for the discrete periodic curvelet transform can also be found in [32]. The wave atom transform will be shortly considered at the end of this section.

For defining a discrete regularization $\mathbf{P}_\sigma : \mathbb{R}^D \rightarrow \mathbb{R}^D$ we use the (matrix) soft threshold function $\Theta = \Theta_s$ or the garrote threshold $\Theta = \Theta_g$. Then $\mathbf{P}_\sigma \mathbf{u}^j = \tilde{\mathbf{u}}^j = \tilde{\mathbf{M}}^T \Theta(\mathbf{M}\mathbf{u}^j) \mathbf{M}\mathbf{u}^j$ depends continuously on \mathbf{u}^j , and we even have

$$\|\mathbf{P}_\sigma \mathbf{u} - \mathbf{P}_\sigma \mathbf{v}\|_2 \leq C \|\mathbf{u} - \mathbf{v}\|_2$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^D$ (see Section 4). This property will be an important assumption for convergence of the digital reaction-diffusion process.

Now we obtain from (3.1) the following iterative filter $H_\lambda : \mathbf{u}^j \rightarrow \mathbf{u}^{j+1}$, which is nonlinear and data-dependent.

For any node $\alpha \in \Omega$ let

$$u_\alpha^{j+1} = H_\lambda(\mathbf{u}^j) = \frac{1}{1 + \lambda} \left(u_\alpha^j + \lambda v_\alpha^0 + \tau \sum_{\beta \sim \alpha} \frac{g(|D_{\alpha,\beta} \tilde{\mathbf{u}}^j|)}{\|e_{\alpha,\beta}\|^2} (u_\beta^j - u_\alpha^j) \right) \quad (3.2)$$

with

$$\tau := \left(\max_{\alpha \in \Omega} \sum_{\beta \sim \alpha} \|e_{\alpha,\beta}\|^{-2} \right)^{-1}. \quad (3.3)$$

Alternatively, the step size τ can be taken as a smaller positive number. As in the continuous model, the diffusivity function $g : \mathbb{R}_+ \rightarrow (0, 1]$ used in the filter is assumed to be a monotone decreasing function with $0 < g(|x|) \leq 1$.

Considering only nodes being directly connected with the node α by an edge in the filter (3.2), we do not need to determine any boundary conditions. The complete algorithm for the proposed denoising filter is now given as follows.

Algorithm

1. Assign a linear order to all nodes $\alpha_1 < \alpha_2 < \dots < \alpha_D$ of Ω and initialize $\mathbf{u}^0 = (u_\alpha^0)_{\alpha \in \Omega}$.
2. Compute the pseudo-observation $\mathbf{v}_0 = (v_\alpha^0)_{\alpha \in \Omega} = \mathbf{S}\mathbf{u}^0$ by wave atom shrinkage.
3. For $j = 1, 2, \dots$
 - Compute $\tilde{\mathbf{u}}^j = \mathbf{P}_\sigma \mathbf{u}^j$ by curvelet shrinkage;
 - For $k = 1, \dots, D$
 - Compute the derivatives $D_{\alpha,\beta} \tilde{\mathbf{u}}^j$;
 - Compute $u_{\alpha_k}^{j+1} = H_\lambda(\mathbf{u}^j)$;
 - End
- End.

The filter H_λ contains the positive regularization parameter λ . As in the continuous model, this parameter is responsible for balancing diffusion and reaction.

Remark. For the special case $P_\sigma = Id$ in (3.1) leading to $\mathbf{u}^j = \tilde{\mathbf{u}}^j$ in (3.2), and $g(|x|)$ being a binary diffusion (i.e., $g(|x|) = 1$ if $|x| < \sigma$ and $g(|x|) = 0$ if $|x| \geq \sigma$ for a given threshold σ), equation (3.2) can for $\lambda = 0$ be reduced to the iterative nonlinear scheme proposed in [41], where the authors have demonstrated equivalence between the nonlinear scheme, adaptive smoothing, bilateral filtering, and translation-invariant Haar wavelet shrinkage.

Wave atom transform

In the following we shortly summarize the wave atom transform as recently suggested in [21]. See also [47] for a very related approach.

We consider a one-dimensional family of wave packets $\psi_{m,n}^j(x)$, $j \geq 0$, $m \geq 0$, $n \in \mathbb{N}$, centered in frequency around $\pm\omega_{j,m} = \pm\pi 2^j m$ with $c_1 2^j \leq m \leq c_2 2^j$ (where $c_1 < c_2$ are suitable positive constants) and centered in space around $x_{j,n} = 2^{-j} n$. For that purpose, let g be a real valued C^∞ bump function with compact support in $[-7\pi/6, 5\pi/6]$ such that for $|\omega| \leq \pi/3$,

$$g\left(\frac{\pi}{2} - \omega\right)^2 + g\left(\frac{\pi}{2} + \omega\right)^2 = 1, \quad g\left(-\frac{\pi}{2} - 2\omega\right) = g\left(\frac{\pi}{2} + \omega\right).$$

Then the function $\hat{\psi}_m^0$ is determined by the formula

$$\hat{\psi}_m^0(\omega) := e^{-i\omega/2} \left[e^{i\alpha_m} g\left(\epsilon_m\left(\omega - \pi\left(m + \frac{1}{2}\right)\right)\right) + e^{-i\alpha_m} g\left(\epsilon_{m+1}\left(\omega + \pi\left(m + \frac{1}{2}\right)\right)\right) \right],$$

where $\epsilon_m = (-1)^m$ and $\alpha_m = \frac{\pi}{2}\left(m + \frac{1}{2}\right)$. The properties of g have to ensure that

$$\sum_{m=0}^{\infty} |\hat{\psi}_m^0(\omega)|^2 = 1.$$

Then the translates $\{\psi_m^0(\cdot - n)\}$ form an orthonormal basis of $L^2(\mathbb{R})$. Introducing the basis functions

$$\psi_{m,n}^j(x) = \psi_m^j(x - 2^{-j}n) = 2^{j/2} \psi_m^0(2^j x - n),$$

the transform $WA : L^2(\mathbb{R}) \rightarrow l^2(\mathbb{Z})$ maps a function u onto a sequence of wave atom coefficients

$$c_{j,m,n} = \int_{-\infty}^{\infty} u(x) \psi_{m,n}^j(x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i2^{-j}n\omega} \overline{\hat{\psi}_m^j(\omega)} \hat{u}(\omega) d\omega.$$

In the two-dimensional case let $\mu = (j, \mathbf{m}, \mathbf{n})$ where $\mathbf{m} = (m_1, m_2)$ and $\mathbf{n} = (n_1, n_2)$. We consider

$$\varphi_\mu^+(x_1, x_2) := \psi_{m_1, n_1}^j(x_1) \psi_{m_2, n_2}^j(x_2)$$

and the Hilbert transformed wavelet packets,

$$\varphi_\mu^-(x_1, x_2) := H\psi_{m_1, n_1}^j(x_1) H\psi_{m_2, n_2}^j(x_2),$$

where for a decomposition $\hat{\psi}_{m,n}^j(\omega) = \hat{\psi}_{m,n,+}^j(\omega) + i\hat{\psi}_{m,n,-}^j(\omega)$ with $\hat{\psi}_{m,n,-}^j(\omega) = \overline{\hat{\psi}_{m,n,+}^j(\omega)}$ the Hilbert transform is defined by

$$\widehat{H}\psi_{m,n}^j(\omega) = -i\hat{\psi}_{m,n,+}^j(\omega) + i\hat{\psi}_{m,n,-}^j(\omega).$$

(Note that the above decomposition of $\hat{\psi}_{m,n}$ is possible since $\psi_{m,n}$ is real-valued.) A recombination

$$\varphi_\mu^{(1)} = \frac{\varphi_\mu^+ + \varphi_\mu^-}{2}, \quad \varphi_\mu^{(2)} = \frac{\varphi_\mu^- + \varphi_\mu^+}{2}$$

provides basis functions with two bumps in the frequency plane being symmetric with respect to the origin. Together, $\varphi_\mu^{(1)}$ and $\varphi_\mu^{(2)}$ form a wave atom frame, and the wave atom coefficients $c_\mu^{(1)}, c_\mu^{(2)}$ are the scalar products of u with $\varphi_\mu^{(1)}$ and $\varphi_\mu^{(2)}$.

In [21], a discretization of this transform is described for the one-dimensional case as well as an extension to two dimensions. The algorithm is based on the fast Fourier transform and a wrapping trick. For implementation software we refer to the homepage <http://www.waveatom.org/software.html> due to Demanet and Ying.

The wave atom shrinkage can be formulated as $u_c = \sum_{\mu} \theta_h(c_{\mu}^{(1)}(u)) \varphi_{\mu}^{(1)} + \theta_h(c_{\mu}^{(2)}(u)) \varphi_{\mu}^{(2)}$, where $\theta_h(x)$ is the hard threshold function mentioned above.

4 Convergence of the discrete method

For showing some properties of the filtering process (3.2) we first derive a matrix-vector representation of the iteration scheme.

Let $\mathbf{u}^j = (u_1^j, \dots, u_D^j)^T$, $\tilde{\mathbf{u}}^j = \mathbf{P}_{\sigma} \mathbf{u}^j = (\tilde{u}_1^j, \dots, \tilde{u}_D^j)^T$, $j = 0, 1, 2, \dots$ and $\mathbf{v}^0 = \mathbf{S} \mathbf{u}^0 = (v_1^0, \dots, v_D^0)^T$. Then the iteration (3.2) can be written in the form

$$\mathbf{u}^{j+1} = \frac{1}{1+\lambda} (\mathbf{u}^j + \lambda \mathbf{v}^0 + \tau \mathbf{G}^j \mathbf{u}^j),$$

where $\mathbf{G}^j = \mathbf{G}(\tilde{\mathbf{u}}^j) = (G_{\alpha,\beta}^j)_{\alpha,\beta=1}^D$ is a sparse matrix depending on $\tilde{\mathbf{u}}^j$ and on the graph given by

$$G_{\alpha,\beta}^j := \begin{cases} -\sum_{\gamma \sim \alpha} \frac{g(|D_{\alpha,\gamma} \tilde{u}^j|)}{\|e_{\alpha,\gamma}\|^2} & \text{for } \beta = \alpha, \\ \frac{g(|D_{\alpha,\beta} \tilde{u}^j|)}{\|e_{\alpha,\beta}\|^2} & \text{for } \beta \sim \alpha, \\ 0 & \text{for } \beta \not\sim \alpha. \end{cases} \quad (4.1)$$

Hence, for the graph on the right hand side of Figure 2, \mathbf{G}^j has at most 9 entries per row. Introducing the matrix

$$\mathbf{A}^j = \mathbf{A}(\tilde{\mathbf{u}}^j) := \mathbf{I} + \tau \mathbf{G}^j, \quad (4.2)$$

where \mathbf{I} denotes the identity matrix of size D , the iteration process reads

$$\mathbf{u}^{j+1} = \frac{1}{1+\lambda} (\mathbf{A}^j \mathbf{u}^j + \lambda \mathbf{v}^0). \quad (4.3)$$

Before studying the convergence of this scheme, we need to consider the properties of the iteration matrix $\mathbf{A}^j = \mathbf{A}(\tilde{\mathbf{u}}^j)$ more closely.

Lemma 4.1 *Let the diffusivity function g satisfy $0 < g(|s|) \leq 1$ for $s \in \mathbb{R}$. Then the iteration matrix \mathbf{A}^j given in (4.2) satisfies the following properties for all $j = 0, 1, 2, \dots$*

1. \mathbf{A}^j is symmetric, i.e., $\mathbf{A}^j = (\mathbf{A}^j)^T$.
2. With $\mathbf{1} := (1, 1, \dots, 1)^T \in \mathbb{R}^d$ we have $\mathbf{A}^j \mathbf{1} = \mathbf{1}$.
3. We have $\mathbf{A}^j \geq \mathbf{0}$, i.e., all entries of \mathbf{A}^j are non-negative. Moreover for the row sum norm and the spectral norm we have $\|\mathbf{A}^j\|_{\infty} = \|\mathbf{A}^j\|_2 = 1$.

Proof. Since $\alpha \sim \beta$ implies $\beta \sim \alpha$, the matrix \mathbf{G}^j and hence also \mathbf{A}^j is symmetric, i.e., $\mathbf{A}^j = (\mathbf{A}^j)^T$. Further, by definition of \mathbf{G}^j we have $\mathbf{G}^j \mathbf{1} = \mathbf{0}$ and hence $\mathbf{A}^j \mathbf{1} = \mathbf{1}$, i.e., 1 is an eigenvalue of \mathbf{A}^j . With τ given in (3.3), we find

$$\tau \sum_{\gamma \sim \alpha} \frac{g(|D_{\alpha,\gamma} \tilde{u}^j|)}{\|e_{\alpha,\gamma}\|^2} \leq \tau \sum_{\gamma \sim \alpha} \frac{1}{\|e_{\alpha,\gamma}\|^2} = 1$$

for all $\alpha \in \Omega$, such that the diagonal entries of \mathbf{A}^j are non-negative, and hence $\mathbf{A}^j \geq \mathbf{0}$. This together with $\mathbf{A}^j \mathbf{1} = \mathbf{1}$ implies for the row sum norm $\|\mathbf{A}^j\|_{\infty} = 1$ and we conclude that the spectral radius of \mathbf{A}^j is 1. \square

Corollary 4.2 *Let the diffusivity function g satisfy $0 < g(|s|) \leq 1$ for $s \in \mathbb{R}$. Then the vectors $\mathbf{u}^j = (u_\alpha^j)_{\alpha \in \Omega}$ obtained by the iteration scheme (3.2) (resp. (4.3)) satisfy*

$$\min_{\alpha \in \Omega} \{u_\alpha^0, v_\alpha^0\} \leq u_\beta^j \leq \max_{\alpha \in \Omega} \{u_\alpha^0, v_\alpha^0\} \quad (4.4)$$

for all $\beta \in \Omega$ and all $j = 1, 2, \dots$ as well as

$$\|\mathbf{u}^j\|_2 = \left(\sum_{\alpha \in \Omega} (u_\alpha^j)^2 \right)^{1/2} \leq \max\{\|\mathbf{u}^0\|_2, \|\mathbf{v}^0\|_2\} \quad \forall j = 1, 2, \dots$$

Further, assuming that the transform S is invariant with respect to the mean value μ of u^0 , i.e., $\mathbf{1}^T \mathbf{u}^0 = \mathbf{1}^T \mathbf{v}^0 = \mu D$, we obtain

$$\frac{1}{D} \sum_{\alpha \in \Omega} u_\alpha^j = \mu,$$

i.e., the mean value of u^j remains to be unchanged during the iteration process.

Proof. Since $\mathbf{A}^j \geq \mathbf{0}$ and $\|\mathbf{A}^j\|_\infty = 1$, the application of \mathbf{A}^j to \mathbf{u}^j is a smoothing procedure, where the elements of $\mathbf{w}^{j+1} = (w_\beta^{j+1})_{\beta \in \Omega} = \mathbf{A}^j \mathbf{u}^j$ are convex linear combinations of entries in \mathbf{u}^j . Hence $\min_{\alpha \in \Omega} u_\alpha^j \leq w_\beta^{j+1} \leq \max_{\alpha \in \Omega} u_\alpha^j$ for all $\beta \in \Omega$ and the iteration scheme (4.3) implies

$$\frac{1}{1+\lambda} (\min_{\alpha \in \Omega} u_\alpha^j + \lambda \min_{\alpha \in \Omega} v_\alpha^0) \leq u_\beta^{j+1} \leq \frac{1}{1+\lambda} (\max_{\alpha \in \Omega} u_\alpha^j + \lambda \max_{\alpha \in \Omega} v_\alpha^0).$$

The assertion (4.4) follows now by induction. Further, by $\|\mathbf{A}^j\|_2 = 1$ we also have

$$\|\mathbf{u}^{j+1}\|_2 \leq \frac{1}{1+\lambda} (\|\mathbf{A}^j \mathbf{u}^j\|_2 + \lambda \|\mathbf{v}^0\|_2) \leq \frac{1}{1+\lambda} (\|\mathbf{u}^j\|_2 + \lambda \|\mathbf{v}^0\|_2) \leq \max\{\|\mathbf{u}^0\|_2, \|\mathbf{v}^0\|_2\}.$$

The conservation of the average value follows by an induction argument from

$$\mathbf{1}^T \mathbf{u}^{j+1} = \frac{1}{1+\lambda} \mathbf{1}^T (\mathbf{A}^j \mathbf{u}^j + \lambda \mathbf{v}^0) = \frac{1}{1+\lambda} (\mathbf{1}^T \mathbf{u}^j + \lambda \mathbf{1}^T \mathbf{v}^0) = \mathbf{1}^T \mathbf{u}^j,$$

where we have used that $\mathbf{1}^T \mathbf{A}^j = \mathbf{1}^T$. □

Remarks. The assertion $\mathbf{1}^T \mathbf{u}^0 = \mathbf{1}^T \mathbf{v}^0 = \mathbf{1}^T \mathbf{S} \mathbf{u}^0$ in Lemma 4.1 is usually satisfied for wavelet and curvelet shrinkage operators, assumed that the shrinkage procedure is only applied to the high pass part of the wavelet filter bank resp. not to the low frequency part of the curvelet transformed signal.

If the operator \mathbf{S} in the reaction-diffusion model is a wavelet shrinkage operator based on an orthogonal wavelet transform, then $\|\mathbf{S} \mathbf{u}^0\|_2 \leq \|\mathbf{u}^0\|_2$ since the application of the shrinkage operator does not increase the norm of a vector. It simply follows that $\|\mathbf{v}^0\|_2 \leq \|\mathbf{u}^0\|_2$ and the assertions of Corollary 4.2 can be suitably simplified.

For the special case $\lambda = 0$ we especially find

$$\min_{\alpha \in \Omega} u_\alpha^0 \leq u_\beta^j \leq \max_{\alpha \in \Omega} u_\alpha^0 \quad \forall \beta \in \Omega, j = 1, 2, \dots,$$

and

$$\|\mathbf{u}^j\|_2 \leq \|\mathbf{u}^0\|_2, \quad \mathbf{1}^T \mathbf{u}^j = \mu D \quad \forall j = 1, 2, \dots$$

Lemma 4.3 *Let \mathbf{A}^j , $j \in \{0, 1, 2, \dots\}$, be the iteration matrix given in (4.2). Further, assume that the graph (Ω, E) has boundary elements, i.e., not all nodes $\alpha \in \Omega$ are endpoints of the same number of edges. Then, for diffusivities g which are decreasing with $g(0) = 1$ and satisfying $0 < g(|s|) \leq 1$, the eigenvalue 1 of \mathbf{A}^j is simple and there exists an $\epsilon > 0$ being independent of j such that all further eigenvalues of \mathbf{A}^j lie inside the interval $(-1 + \epsilon, 1 - \epsilon)$.*

The proof of this Lemma can be directly obtained from Lemma 3 in [40]. While in [40], the iteration matrix is defined with a slightly different matrix \mathbf{G}^j , all properties of \mathbf{A}^j remain to be satisfied, such that the result of Lemma 4.3 holds.

We are now ready to consider the convergence of the new discrete reaction-diffusion scheme (3.2). First, we shall consider convergence of the iteration process (3.2) for $\lambda = 0$.

Theorem 4.4 *Let $(\mathbf{u}^j)_{j \geq 0}$ be the sequence of vectors obtained by the iteration process (3.2) (resp. (4.3)) applied to a starting vector \mathbf{u}^0 for $\lambda = 0$. Then $(\mathbf{u}^j)_{j \geq 0}$ converges to the mean value $\mu \mathbf{1}$ with $\mu := \frac{1}{D} \sum_{\alpha \in \Omega} u_\alpha^0$.*

Proof. The proof is analogous to that of Theorem 4.4 in [40]. Let us consider $\mathbf{r}^j := \mathbf{u}^j - \mu \mathbf{1}$. Then we have by Corollary 4.2 and the corresponding remark $\mathbf{r}^{j+1} = \mathbf{A}^j \mathbf{r}^j$. Since \mathbf{r}^j is orthogonal to the eigenvector $\mathbf{1}$ of \mathbf{A}^j corresponding to the eigenvalue 1, it follows for the Euclidean norm by Lemma 4.3 that

$$\|\mathbf{r}^{j+1}\|_2 = \|\mathbf{A}^j \mathbf{r}^j\|_2 \leq (1 - \epsilon) \|\mathbf{r}^j\|_2.$$

Thus $\lim_{j \rightarrow \infty} \mathbf{r}^j = \mathbf{0}$ and the convergence $\lim_{j \rightarrow \infty} \mathbf{u}^j = \mu \mathbf{1}$ follows. \square

Remark. Observe that the above convergence result does not depend on the regularization operator P_σ . In fact this result is also true without regularization, i.e., for $P_\sigma = Id$, see [40]. The convergence of the Perona-Malik process ($\lambda = 0$) in the discrete case has been proved already by Weickert (see [48], pp. 97). In [41], a similar iteration process for discrete diffusion has been considered with a special discontinuous diffusivity function g . In this case convergence to a piecewise constant image has been shown.

Further, for sufficiently large λ we find

Theorem 4.5 *Let $(\mathbf{u}^j)_{j \geq 0}$ be the sequence of vectors obtained by the iteration process (3.2) (resp. (4.3)) applied to a starting vector \mathbf{u}^0 . Let $\mathbf{A}(\tilde{\mathbf{u}}) = \mathbf{A}(P_\sigma \mathbf{u})$ denote the iteration matrix as in (4.2) for the vector $\tilde{\mathbf{u}} = P_\sigma \mathbf{u} \in \mathbb{R}^D$. Further, let the Lipschitz condition*

$$\|\mathbf{A}(\tilde{\mathbf{u}}) - \mathbf{A}(\tilde{\mathbf{v}})\|_2 \leq C \|\mathbf{u} - \mathbf{v}\|_2$$

be satisfied for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^D$ with a constant $C > 0$. Then, the sequence $(\mathbf{u}^j)_{j \geq 0}$ converges for all $\lambda > C(\max\{\|\mathbf{u}^0\|_2, \|\mathbf{v}^0\|_2\})$.

Proof. As in [40] one can simply show that $(\mathbf{u}^j)_{j \geq 0}$ is a Cauchy sequence in the $\|\cdot\|_2$ -norm. \square

Remarks. 1. Observe that a Lipschitz condition as in Theorem 4.5 is satisfied for continuous diffusivity functions g and a continuous operator P_σ . Indeed, most diffusivity functions used in image processing are continuous as the Perona-Malik diffusivity and the

Charbonnier diffusivity. The Gaussian regularization and the curvelet regularization with soft threshold or garrote threshold also satisfy the assumption. In particular, using soft threshold we find with the matrix representation of the discrete operator $\mathbf{P}_\sigma : \mathbb{R}^D \rightarrow \mathbb{R}^D$ the estimate

$$\begin{aligned} \|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|_2 &= \|\mathbf{P}_\sigma \mathbf{u} - \mathbf{P}_\sigma \mathbf{v}\| \\ &= \|\tilde{\mathbf{M}}^T \Theta_s(\mathbf{M}\mathbf{u}) \mathbf{M}\mathbf{u} - \tilde{\mathbf{M}}^T \Theta_s(\mathbf{M}\mathbf{v}) \mathbf{M}\mathbf{v}\|_2 \\ &\leq \|\tilde{\mathbf{M}}^T\|_2 \left(\|\Theta_s(\mathbf{M}\mathbf{u})(\mathbf{M}\mathbf{u} - \mathbf{M}\mathbf{v})\|_2 + \|\Theta_s(\mathbf{M}\mathbf{u}) - \Theta_s(\mathbf{M}\mathbf{v})\|_2 \|\mathbf{M}\mathbf{v}\|_2 \right). \end{aligned}$$

Here, \mathbf{M} and $\tilde{\mathbf{M}}$ are $\tilde{D} \times D$ matrices and $\|\mathbf{M}\|_2^2 := \|\mathbf{M}\mathbf{M}^T\|_2$. With $\hat{\mathbf{u}} = \mathbf{M}\mathbf{u} \in \mathbb{R}^{\tilde{D}}$ and $\hat{\mathbf{v}} = \mathbf{M}\mathbf{v} \in \mathbb{R}^{\tilde{D}}$ we obtain

$$\Theta_s(\hat{\mathbf{u}}) = \text{diag} \left(\left(\max\left\{ \frac{|\hat{u}_k| - \sigma}{|\hat{u}_k|}, 0 \right\} \right)_{k=1}^{\tilde{D}} \right)$$

with the convention $\frac{-\sigma}{0} = -\infty$. Thus, $\|\Theta_s(\hat{\mathbf{u}})\| < 1$ and a simple case study yields

$$\begin{aligned} \|\Theta_s(\hat{\mathbf{u}}) - \Theta_s(\hat{\mathbf{v}})\|_2 &= \left\| \text{diag} \left(\left(\max\left\{ \frac{|\hat{u}_k| - \sigma}{|\hat{u}_k|}, 0 \right\} - \max\left\{ \frac{|\hat{v}_k| - \sigma}{|\hat{v}_k|}, 0 \right\} \right)_{k=1}^{\tilde{D}} \right) \right\|_2 \\ &\leq \left\| \text{diag} \left(\left(\frac{|\hat{u}_k - \hat{v}_k|}{\sigma} \right)_{k=1}^{\tilde{D}} \right) \right\|_2 \leq \frac{1}{\sigma} \max\{|\hat{u}_k - \hat{v}_k|, k = 1, \dots, \tilde{D}\} \\ &\leq \frac{1}{\sigma} \|\hat{\mathbf{u}} - \hat{\mathbf{v}}\|_2 \leq \frac{1}{\sigma} \|\mathbf{M}\|_2 \|\mathbf{u} - \mathbf{v}\|_2. \end{aligned}$$

Hence, we obtain the final estimate

$$\|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|_2 = \|\mathbf{P}_\sigma \mathbf{u} - \mathbf{P}_\sigma \mathbf{v}\|_2 \leq \|\tilde{\mathbf{M}}^T\|_2 \|\mathbf{M}\|_2 \left(1 + \frac{1}{\sigma} \|\mathbf{M}\|_2 \|\mathbf{v}\|_2 \right) \|\mathbf{u} - \mathbf{v}\|_2.$$

2. The above result on the convergence of the iteration scheme (3.2) for sufficiently large λ is in accordance with the results of Nordström for another discretization of the Euler equations in [36].

As far as we know, there is no result in the literature up to now, showing the convergence of a reaction-diffusion process like in (3.2) for small λ . The numerical results suggest convergence of this scheme also for small $\lambda > 0$. Next we shall prove convergence of the iteration scheme (3.2) in a generalized sense.

Theorem 4.6 *Let $(\mathbf{u}^j)_{j \geq 0}$ be the sequence of vectors obtained by the iteration process (3.2) (resp. (4.3)) applied to a starting vector \mathbf{u}^0 . Let $\mathbf{A}(\tilde{\mathbf{u}}) = \mathbf{A}(P_\sigma \mathbf{u})$ denote the iteration matrix as in (4.2) for the vector $\tilde{\mathbf{u}} = P_\sigma \mathbf{u} \in \mathbb{R}^D$. Further, let the Lipschitz condition*

$$\|\mathbf{A}(\tilde{\mathbf{u}}) - \mathbf{A}(\tilde{\mathbf{v}})\|_2 \leq C \|\mathbf{u} - \mathbf{v}\|_2$$

be satisfied for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^D$ with a constant $C > 0$. Assume that the sequence $(\mathbf{u}^j)_{j \geq 0}$ has at most a finite number of accumulation points. Then, the sequence $(\mathbf{u}^j)_{j \geq 0}$ is Cesàro-convergent for all $\lambda \geq 0$, i.e., the sequence $(\frac{1}{n} \sum_{j=0}^{n-1} \mathbf{u}^j)_{n \geq 0}$ converges.

Proof. By Corollary 4.2 the sequence $(\mathbf{u}^j)_{j \geq 0}$ is bounded. Hence, there exists a partial sequence $(\mathbf{u}^{j_k})_{k \geq 0}$ that converges to a bounded limit. Then we find for every $\epsilon > 0$ a constant $M_\epsilon \in \mathbb{N}$ such that

$$\|\mathbf{u}^{j_k} - \mathbf{u}^{j_{k+m}}\|_2 < \epsilon \quad \text{for all } k \geq M_\epsilon, m \geq 1.$$

We show now that the sequence $(\mathbf{u}^{j_k+1})_{k \geq 0}$ is also a Cauchy sequence. Using Lemma 4.1, Corollary 4.2 and the assumption of the theorem we have

$$\begin{aligned} \|\mathbf{u}^{j_k+1} - \mathbf{u}^{j_{k+m}+1}\|_2 &= \frac{1}{1+\lambda} \|\mathbf{A}^{j_k} \mathbf{u}^{j_k} - \mathbf{A}^{j_{k+m}} \mathbf{u}^{j_{k+m}}\|_2 \\ &\leq \frac{1}{1+\lambda} (\|\mathbf{A}^{j_k} (\mathbf{u}^{j_k} - \mathbf{u}^{j_{k+m}})\|_2 + \|(\mathbf{A}^{j_k} - \mathbf{A}^{j_{k+m}}) \mathbf{u}^{j_{k+m}}\|_2) \\ &\leq \frac{1}{1+\lambda} (\|\mathbf{u}^{j_k} - \mathbf{u}^{j_{k+m}}\|_2 + C \|\mathbf{u}^{j_k} - \mathbf{u}^{j_{k+m}}\|_2 \|\mathbf{u}^{j_{k+m}}\|_2) \\ &\leq \frac{1+CR}{1+\lambda} \epsilon, \end{aligned}$$

with $R := \max\{\|\mathbf{u}^0\|_2, \|\mathbf{v}^0\|_2\}$. Hence, the sequence $(\mathbf{u}^{j_k+1})_{k \geq 0}$ also converges. This argument can successively be applied. Let now $\tilde{\mathbf{u}}^n$ be the limit of $(\mathbf{u}^{j_k+n})_{k \geq 0}$ for $n = 0, 1, \dots$. Hence there exist matrices $\tilde{\mathbf{A}}^n$ with $\lim_{k \rightarrow \infty} \mathbf{A}^{j_k+n} = \tilde{\mathbf{A}}^n$ and we find

$$\begin{aligned} \tilde{\mathbf{u}}^n &= \lim_{k \rightarrow \infty} \mathbf{u}^{j_k+n} = \lim_{k \rightarrow \infty} \frac{1}{\lambda+1} (\mathbf{A}^{j_k+n-1} \mathbf{u}^{j_k+n-1} + \lambda \mathbf{v}^0) \\ &= \frac{1}{\lambda+1} (\tilde{\mathbf{A}}^{n-1} \tilde{\mathbf{u}}^{n-1} + \lambda \mathbf{v}^0) \end{aligned}$$

for all $n \in \mathbb{N}$, i.e., the accumulation points $\tilde{\mathbf{u}}_n$ also satisfy the iteration equation. Since the number of accumulation points is finite, there exists a smallest number $r \in \mathbb{N}$ such that for some $s \in \mathbb{N}$ with $s < r$ the two sequences $(\mathbf{u}^{j_k+s})_{k \geq 0}$ and $(\mathbf{u}^{j_k+r})_{k \geq 0}$ have the same limit $\tilde{\mathbf{u}}^s$. Hence we find

$$\lim_{k \rightarrow \infty} \mathbf{u}^{j_k+r+1} = \lim_{k \rightarrow \infty} \frac{1}{\lambda+1} (\mathbf{A}^{j_k+r} \mathbf{u}^{j_k+r} + \lambda \mathbf{v}^0) = \frac{1}{\lambda+1} (\tilde{\mathbf{A}}^s \tilde{\mathbf{u}}^s + \lambda \mathbf{v}^0) = \tilde{\mathbf{u}}^{s+1}.$$

Repeating this argument, we obtain that there are only the pairwise different accumulation points $\tilde{\mathbf{u}}^s, \dots, \tilde{\mathbf{u}}^{r-1}$. Hence $s = 0$ and $(\mathbf{u}^{k_r+\nu})_{\nu \geq 0}$ converges for $\nu = 0, 1, \dots, r-1$. We conclude that the new sequence

$$\mathbf{s}_k := \frac{1}{r} \sum_{\nu=0}^{r-1} \mathbf{u}^{rk+\nu}$$

also converges. Finally, applying the Theorem of Stolz-Cesàro, the Cesàro convergence of $(\mathbf{u}^j)_{j \geq 0}$ follows from

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=0}^{n-1} \mathbf{u}^\nu &= \lim_{n \rightarrow \infty} \left(\frac{r}{n} \sum_{k=1}^{\lfloor n/r \rfloor - 1} \mathbf{s}_k + \frac{1}{n} \sum_{\nu=\lfloor n/r \rfloor r}^{n-1} \mathbf{u}^\nu \right) \\ &= \lim_{n \rightarrow \infty} \frac{r}{n} \sum_{k=1}^{\lfloor n/r \rfloor - 1} \mathbf{s}_k + \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=\lfloor n/r \rfloor r}^{n-1} \mathbf{u}^\nu \\ &= \lim_{n \rightarrow \infty} \mathbf{s}_{\lfloor n/r \rfloor - 1} = \lim_{n \rightarrow \infty} \mathbf{s}_n. \end{aligned}$$

In the special case that we find partial sequences $(\mathbf{u}^{j_k})_{k \geq 0}$ and $(\mathbf{u}^{j_{k+1}})_{k \geq 0}$ with the same limit $\tilde{\mathbf{u}}$, the above arguments lead to usual convergence of the complete sequence $(\mathbf{u}^k)_{k \geq 0}$ to $\tilde{\mathbf{u}}$. \square

Finally, we remark that our digital diffusion-reaction type filter can be seen as digital analogon of the Euler equations that are obtained as a steady state of (2.2).

Theorem 4.7 *If the filtering sequence $(\mathbf{u}^j)_{j \geq 0}$ given by the scheme (3.2) converges to a limit vector $\mathbf{u}^* \in \mathbb{R}^D$, then \mathbf{u}^* satisfies*

$$\frac{\tau}{2} \sum_{e \succ \alpha} \frac{\partial}{\partial e} (-g(|D_e \tilde{\mathbf{u}}^*|)) \frac{\partial u^*}{\partial e} |_{\alpha} + \lambda (v_{\alpha}^0 - u_{\alpha}^*) = 0 \quad \forall \alpha \in \Omega.$$

Here, again $e \succ \alpha$ means, that e is an edge with the endpoint α .

The proof of this assertion can be obtained in the same manner as in [40], Theorem 4.6. \square

5 Numerical experiments

Now we display the good performance of our proposed method in comparison to some existing methods. Fig. 3 (a) is a part of noisy Barbara image. Fig. 3 (b) is a denoised result obtained by a tensor-product decimated Db6 wavelet shrinkage with 3-level decomposition. Nonsmooth oscillations along the edges can be seen obviously due to the poor ability of wavelets at presenting line singularities. Fig. 3 (c) and Fig. 3 (d) are obtained using curvelet transform and wave atom transform, respectively. The hard threshold is hired for these two transforms. Although the curvelet and wave transform are effective in recovering edges and textures, they suffer from the pseudo-Gibbs and element-like artifacts yet. Fig. 3 (e) shows the result by classical TV diffusion (without reaction) with step size 0.002 and 25 iterations. Fig. 3 (f) shows the removed components by Fig. 3 (e). Fig. 3 (g) shows the result by our proposed reaction-diffusion filter with Perona-Malik diffusivity ($\kappa = 0.02$), fitting parameter $\lambda = 0.2$, step size 0.05, and 15 iterations, $\sigma = 0.15$ for wave atom hard threshold and a scale-dependent threshold for curvelet shrinkage. Fig. 3 (h) shows the removed components by Fig. 3 (g). It can be seen clearly that the classical TV diffusion loses the textural details while our method keeps the textures and edges well in recovery images when we remove the noise. The parameters that we used for each method have been taken to optimize the SNR and visual quality of each method independently. Fig. 4 (a) and Fig. 4 (b) display the rate of change of SNR for an increasing number of iterations for above TV diffusion and the proposed reaction-diffusion filter, respectively. The horizontal coordinate denotes the number of iterations, and the vertical coordinate denotes the SNR value of recovery images. Our new method can achieve much higher SNR. The parameter λ relates to the balance of diffusion and feedback of reaction. In Fig. 4 (b), we show the SNR vs. iteration for various values of λ . The dashed line, real line, dot-dashed line, and dotted line denote the $\lambda = 0.1, 0.2, 0.3, 0.4$ respectively. Normally, a large λ slows down the smoothing process. For very large λ the iterated scheme yields a sequence of images that approach Su_0 such that the pseudo-Gibbs artifacts can not suppressed effectively.

Furthermore, Figure 5, we compare the proposed the method with existing digital TV diffusion [15] and digital diffusion-reaction type filter [40]. Fig. 5 (left column) shows the results by the digital TV diffusion with $\lambda = 80$, while Fig. 5 (right column) shows the

results by the digital diffusion-reaction type filter with PM diffusivity $\kappa = 0.02$, $\lambda = 0.2$, step size 0.05 and 15 iterations. The upper row displays the denoised results and the lower row displays removed components. Both methods can not preserve the textures well when one kicks out noise.

Fig. 6 shows the performance of our proposed method for images with heavy noise. Textures can be preserved in this case by the method with parameter $\kappa = 0.02$, $\lambda = 0.15$, $\sigma = 0.55$, and step size 0.1.

We also test our method for real images. Figure 7(a) is a measured engineering surface image from industry. Fig. 7 (b) and (c) are denoised results by decimated Db6 wavelet transform and our new reaction-diffusion method, respectively. The wavelet method destroys the structures to some extent, while our method preserves the scratches well in recovery surfaces.

Finally, it should be noted that the graph shown in Figure 2 (right) has been used for spatial discretization of the Equation (2.2), and the practical codes in this paper were developed by MATLAB based on CurveLab package, which is available from www.curvlet.org.

6 Conclusion

In this paper, we proposed a digital reaction-diffusion filter for denoising of textural images by combining new tools from harmonic analysis such as curvelets and wave atoms with a nonlinear diffusion equation. The curvelet shrinkage is used as the regularization of the ill-posed diffusion, and wave atom shrinkage is hired as the pseudo-observation reaction. The convergence of the discrete scheme of the reaction-diffusion process is proved for arbitrary regularization parameter $\lambda \geq 0$. Numerical experiments show the good performance of the proposed method in comparison to some recent techniques.

The general reaction-diffusion method is called image assimilation. It is an ensemble of techniques combining the mathematical information provided by equations and the physical information given by observations in an optimal way, in order to retrieve as best as possible the state variables of a model. The next work is to analyse error influence and the degree of influence of fitting parameter λ and threshold parameter σ in the model.

References

- [1] S. T. Acton, D. Mukherjee, J. Havlicek, A. C. Bovik, Oriented texture completion by AM-FM reaction-diffusion, *IEEE Trans. Image Process.* **10** (6) (2001), 885-896.
- [2] L. Alvarez, J. Esclarin, Image quantization using reaction-diffusion equations, *SIAM J. Appl. Math.* **57** (1) (1997), 153-175.
- [3] G. Aubert, J.-F. Aujol, Modeling very oscillating signals. Application to image processing, *Appl. Math. Optimization* **51** (2) (2005), 163-182.
- [4] J.-F. Aujol, A. Chambolle, Dual norms and image decomposition models, *Int. J. Computer Vision* **63**, (2005), 85-1004.
- [5] C. Barcelos, M. Boaventura, E. Silva, A well-balanced flow equation for noise removal and edge detection, *IEEE Trans. Image Process.* **12** (7) (2003), 751-763.
- [6] T. Berger, J. Stromberg, T. Eltoft, Adaptive regularized constrained least squares image restoration, *IEEE Trans. Image Process.* **8** (1999), 1191-1203.
- [7] A. Buades, B. Coll, J. M. Morel, A review of image denoising algorithms, with a new one, *Multiscale Model. Simul.* **4** (2) (2005), 490-530.

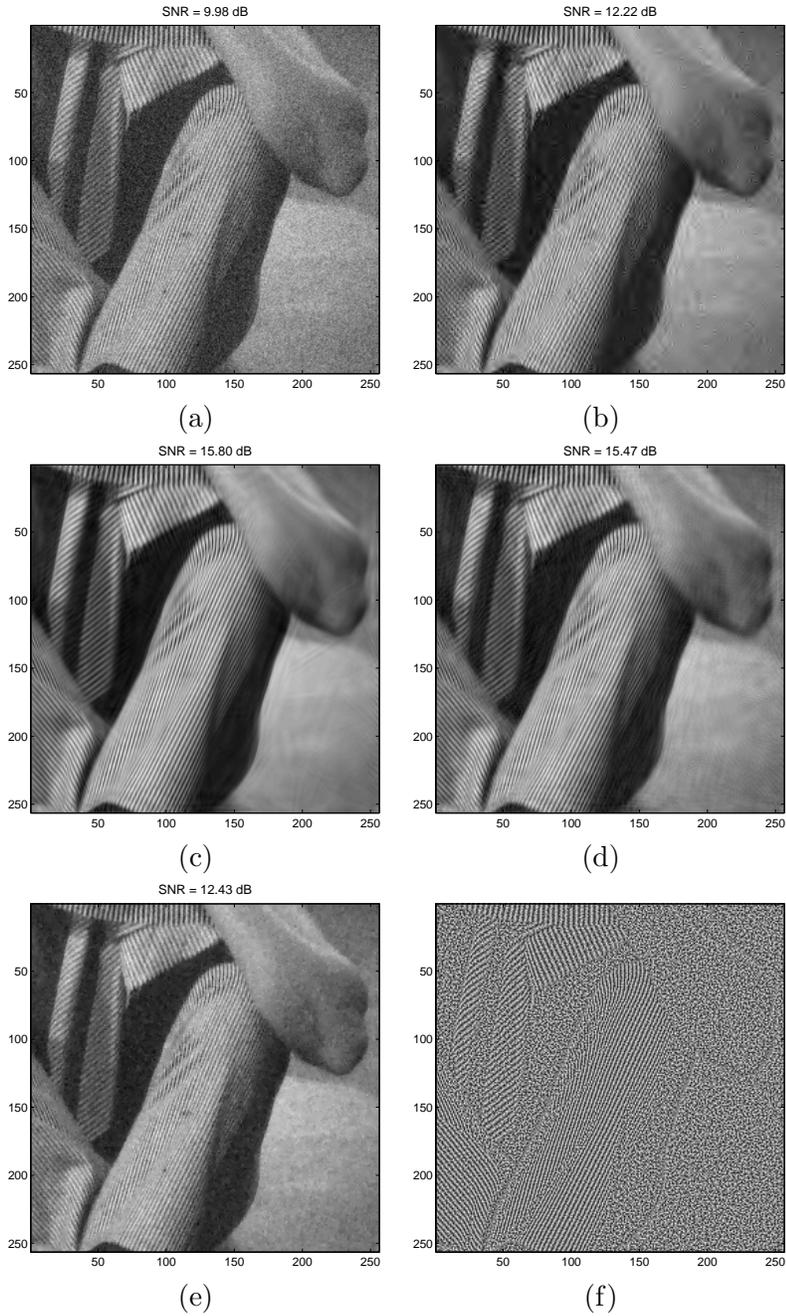


Figure 3: Denoising using different methods. (a) noisy image (SNR=9.98 dB); (b) denoising by wavelets (SNR=12.22 dB); (c) by curvelets (SNR=15.80 dB); (d) by wave atoms (SNR=15.47 dB); (e) by TV diffusion (SNR=12.43 dB); (f) removed components by (e); (g) by the proposed method (SNR=16.86 dB); (h) removed components by (g).

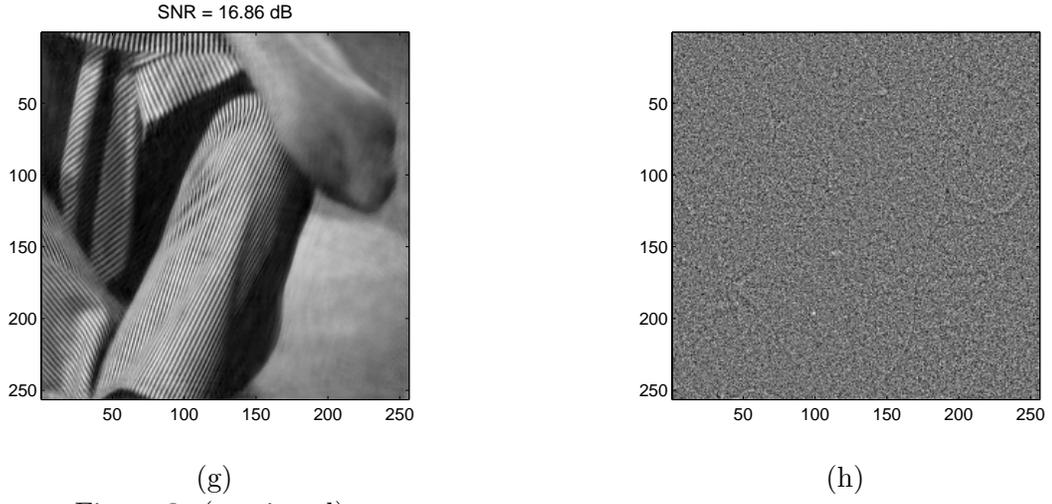
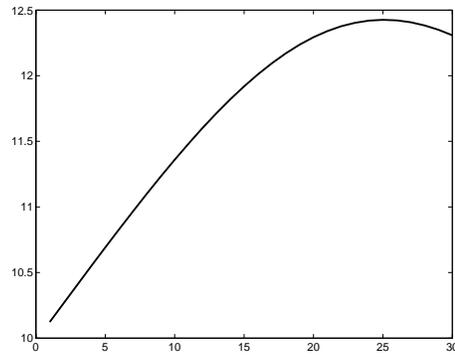
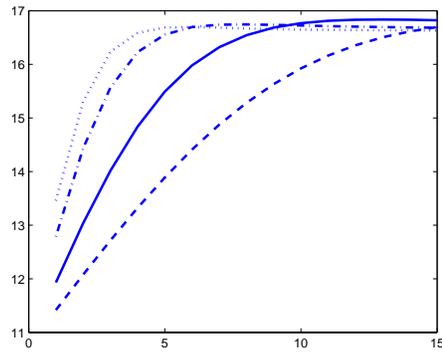


Figure 3: (continued).

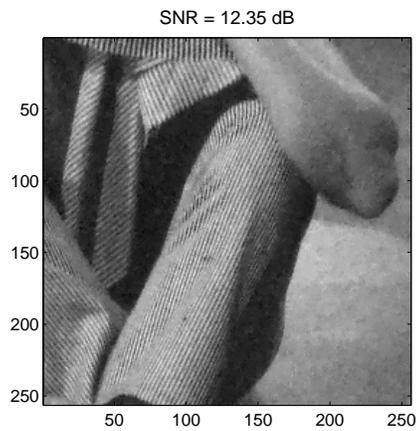


(a)

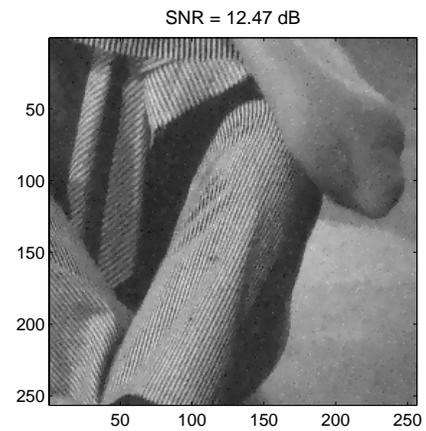


(b)

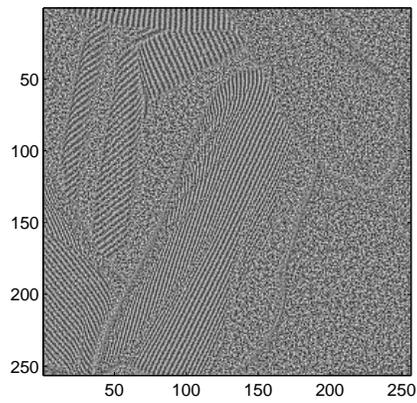
Figure 4: Rate of change of SNR (vertical axis) vs. iteration number (horizontal axis) for classic TV diffusion (a) and the proposed method (b). In (b), the dashed line, real line, dot-dashed line, and dotted line denote for $\lambda = 0.1, 0.2, 0.3, 0.4$ respectively.



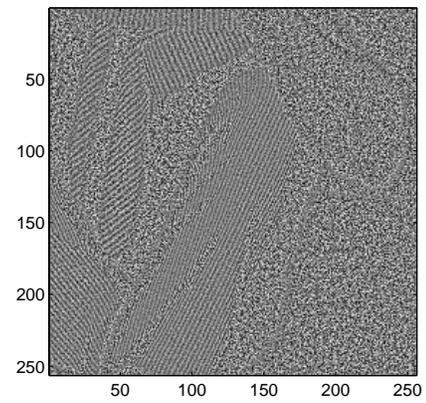
(a)



(b)



(c)



(d)

Figure 5: Denoising by (a) digital TV diffusion [15] (SNR=12.35 dB) and (b) digital diffusion-reaction type filter [40] (SNR=12.47 dB). (c) and (d) are removed components by method (a) and (b).

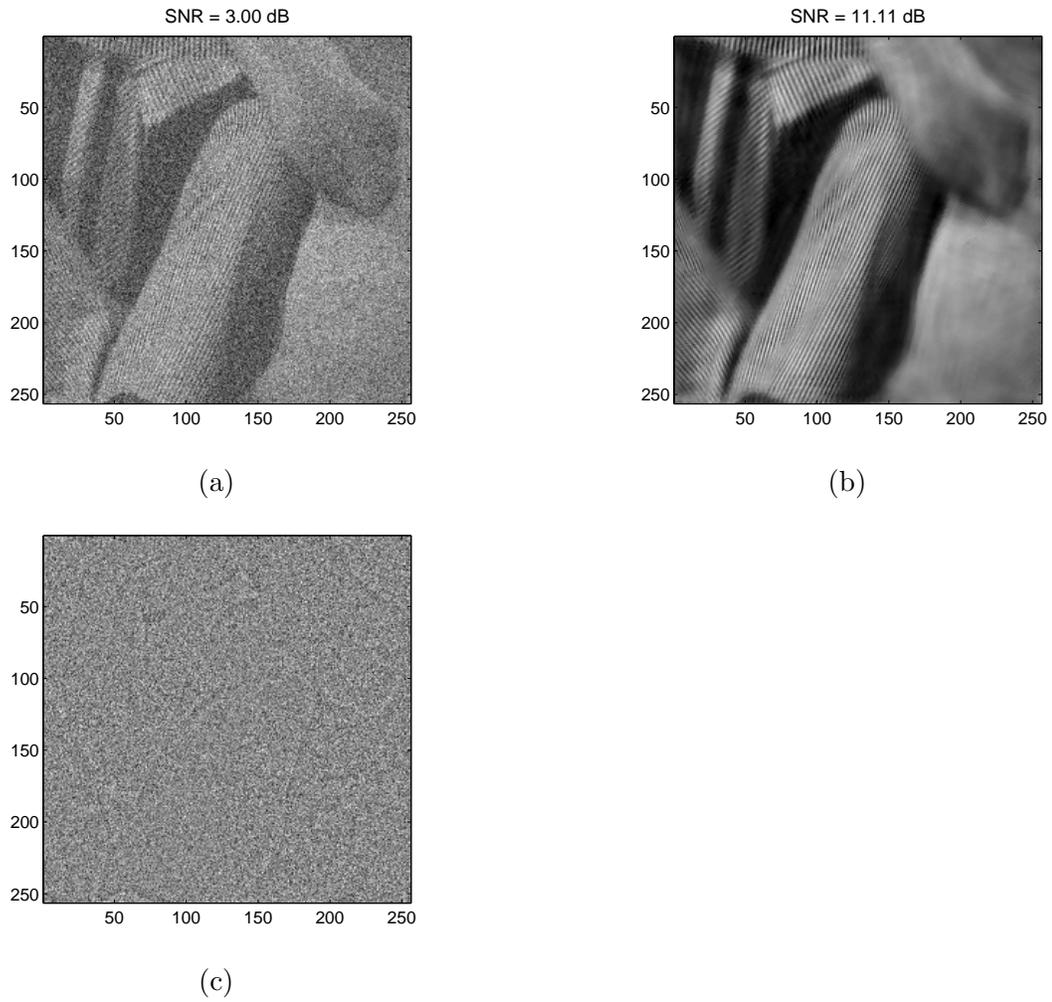


Figure 6: Application of the proposed method for image with heavy noise. (a) noisy image (SNR=3.00 dB); (b) denoised image by the proposed method (SNR=11.11 dB); (c) removed components.

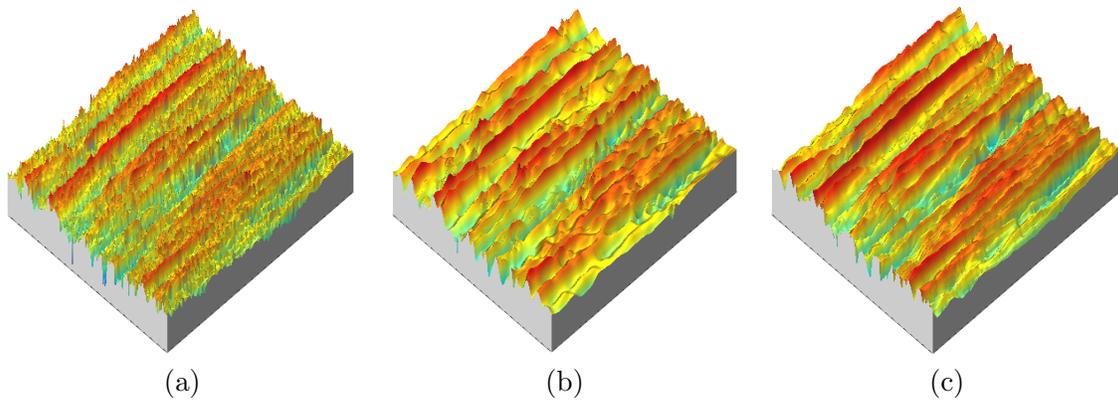


Figure 7: Denoising of a real measured engineering surface (a) by wavelets (b) and the proposed method (c).

- [8] E. J. Candès, D. L. Donoho, Curvelets - a surprisingly effective nonadaptive representation for objects with edges, in *Curves and Surface Fitting: Saint-Malo 1999*, A. Cohen, C. Rabut, L. L. Schumaker (Eds.), Vanderbilt Univ. Press, Nashville, 2000, 105-120.
- [9] E. J. Candès, D. L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise singularities, *Comm. Pure Appl. Math.* **57** (2004), 219-266.
- [10] E. J. Candès, L. Demanet, D. L. Donoho, L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.* **5** (2006), 861-899.
- [11] J. M. Carnicer, W. Dahmen, J. M. Pena, Local Decomposition of Refinable Spaces and Wavelets, *Appl. Comp. Harmon. Anal.* **3** (1996), 127-153.
- [12] F. Catté, P. L. Lions, J. M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* **32** (1992), 1895-1909.
- [13] A. Chambolle, P. L. Lions, Image recovery via total variation minimization and related problems, *Numer. Math.* **76** (1997), 167-188.
- [14] A. Chambolle, R. A. DeVore, N. Lee, B. J. Lucier, Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage, *IEEE Trans. Image Process.* **7** (3) (1998), 319-335.
- [15] T. F. Chan, S. Osher, J. Shen, The digital TV filter and nonlinear denoising, *IEEE Trans. Image Process.* **10** (2001), 231-241.
- [16] Y. Chen, B. Vemuri, L. Wang, Image denoising and segmentation via nonlinear diffusion, *Comput. Math. Appl.* **39** (2000), 131-149.
- [17] R. Coifman, D. Donoho, Translation invariant denoising, In: A. Antoniadis and G. Oppenheim (Eds.), *Wavelets in Statistics*, Springer, New York, 1995, 125-150.
- [18] G. Cottet, L. Germain, Image processing through reaction combined with nonlinear diffusion, *Math. Comput.* **61** (1993), 659-673.
- [19] I. Daubechies, M. Defrise, C. De Mol, An iterative threshold algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.* **57** (2004), 1413-1457.
- [20] I. Daubechies, G. Teschke, Variational image restoration by means of wavelets simultaneous decomposition, deblurring, and denoising, *Appl. Comput. Harmon. Anal.* **19** (2005), 1-16.
- [21] L. Demanet, L. Ying, Wave atoms and sparsity of oscillatory patterns, *Appl. Comput. Harmon. Anal.*, **23** (3) (2007), 368-387.
- [22] S. Esedoglu, An analysis of the Perona-Malik scheme, *Comm. Pure Appl. Math.* **54** (2001), 1442-1487.
- [23] S. Esedoglu, Stability properties of the Perona-Malik scheme, *SIAM J. Numer. Anal.* **44** (3) (2006), 1297-1313.
- [24] G. Gilboa, S. Osher, Nonlocal Linear Image Regularization and Supervised Segmentation, *SIAM Multiscale Model. Simul.* **6**(2) (2007), 595-630.
- [25] J. B. Garnett, T. M. Le, Y. Meyer, L. A. Vese, Image decompositions using bounded variation and generalized homogeneous Besov spaces, *Appl. Comput. Harmon. Anal.* **23** (2007), 25-56.
- [26] S. Kichenassamy, The Perona-Malik paradox, *SIAM J. Appl. Math.* **57** (5) (1997), 1328-1342.
- [27] T. M. Le, L. A. Vese, Image Decomposition using total variation and $div(BMO)$, *Multiscale Model. Simul.* **4** (2) (2005), 390-423.
- [28] W.-Q. Lim, Wavelets with composite dilations, Ph.D. Thesis, Washington University in St. Louis, 2006.

- [29] J. Ma, A. Antoniadis, F.-X. Le Dimet, Curvelet-based snake for multiscale detection and tracking of geophysical fluids, *IEEE Trans. Geosci. Remote Sensing* **44** (12) (2006), 3626-3638.
- [30] J. Ma, Curvelets for surface characterization, *Appl. Phys. Lett.* **90** (2007), 054109: 1-3.
- [31] J. Ma, Image assimilation by geometric wavelet based reaction-diffusion equation, Proceeding of SPIE Wavelet XII, vol. 6701, San Diego, August 26-29, 2007.
- [32] J. Ma, G. Plonka, Combined curvelet shrinkage and nonlinear anisotropic diffusion, *IEEE Trans. Image Process.* **16** (9) (2007), 2198-2206.
- [33] Y. Meyer, Oscillating patterns in image processing and nonlinear evolution equations, University Lecture Series, Vol. 22, Amer. Math. Soc., 2001.
- [34] M. Mignotte, A segmentation-based regularization term for image deconvolution, *IEEE Trans. Image Process.* **15** (2006), 1973-1984.
- [35] M. Nikolova, Local strong homogeneity of a regularized estimator, *SIAM J. Appl. Math.* **61** (2000), 633-658.
- [36] N. Nordström, Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection, *Image Vis. Comput.* **8** (1990), 318-327.
- [37] S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, An iterative regularization method for total variation-based image restoration, *Multiscale Model. Simul.* **4** (2005), 460-489.
- [38] S. Osher, A. Sole, L. Vese, Image decomposition and restoration using total variation minimization and H^{-1} norm, *Multiscale Model. Simul.* **1** (2003), 349-370.
- [39] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **12** (1990), 629-639.
- [40] G. Plonka, A digital diffusion-reaction type filter for nonlinear denoising, *Results in Mathematics* (2008), to appear.
- [41] G. Plonka, J. Ma, Convergence of an iterative nonlinear scheme for denoising of piecewise constant images, *Int. J. Wavelets, Multiresolution and Information Process.* **5** (6) (2007), 975-995.
- [42] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Phys. D* **60** (1992), 259-268.
- [43] J. L. Starck, E. J. Candès, D. L. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11** (2002), 670-684.
- [44] J.L. Starck, M. Elad, D.L. Donoho, Image decomposition via the combination of sparse representations and a variational approach, *IEEE Trans. Image Process.* **14**(10) (2005), 1570-1582.
- [45] F. Torkamani-Azar, K. E. Tait, Image recovery using anisotropic diffusion equation, *IEEE Trans. Image Process.* **5** (1996), 1573-1578.
- [46] L. Vese, S. Osher, Modeling textures with total variation minimization and oscillating patterns in image processing, *J. Sci. Comput.* **19** (2003), 553-572.
- [47] L. Villemoes, Wavelet packets with uniform time-frequency localization, *C. R. Acad. Sci. Paris, Ser. I* **335** (2002), 793-796.
- [48] J. Weickert, Anisotropic Diffusion in Image Processing, Teubner, Stuttgart, 1998.
- [49] J. Weickert, B. M. ter Haar Romeny, M. A. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Trans. Image Process.* **7**(3) (1998), 398-410.
- [50] M. Welk, G. Steidl, J. Weickert, A four-pixel scheme for singular differential equations. In R. Kimmel, N. Sochen, J. Weickert (Eds.), *Scale-Space and PDE Methods in Computer Vision*. LNCS, Springer, Berlin, 610-621 (2005).

- [51] J. Z. Wang, Wavelet oriented anisotropic diffusion in image enhancement, Technical Report, 2004.
- [52] S. Zhu, D. Mumford, Prior learning and Gibbs reaction-diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **19** (11) (1997), 1236-1250.