A digital diffusion-reaction type filter for nonlinear denoising

Gerlind Plonka

Abstract. We propose a new digital diffusion-reaction filter for denoising of digital images and data living on graphs. The construction of this filter is motivated by the regularization and diffusion approach to edge detection by Nordström. Properties and convergence of the iteration scheme are studied and numerical results are given.

§1. Introduction

Image denoising is a field where one is typically interested in removing noise, which may be introduced by the image information process, image recording, image transmission, etc. In a wide variety of applications, the images are discontinuous, and the challenge is to smooth them while preserving their edges and important structures.

Different nonlinear methods have been proposed to tackle this problem including global stochastic methods, adaptive smoothing, wavelet techniques, anisotropic diffusion and variational methods (see e.g. [1,2,4,8, 9,12,15,16]).

Recently, close connections, similarities and relations between these method classes have been examined. New hybrid methods have been proposed which aim to connect the advantageous properties of the different denoising techniques (see e.g. [3,7,10,13,14]).

In this paper, we propose a new nonlinear data-dependent denoising filter, called digital diffusion-reaction filter. It can be seen as an adaption of the digital total variation filter introduced by Chan, Osher and Shen [4]. The TV-filter in [4] can be interpreted as a precise translation of the classical analog TV restoration model invented by Rudin, Osher and Fatemi [12] to the digital case.

Our model can be understood as a digital variant of a diffusion-reaction type equation suggested by Nordström [8]. The Euler equations associated with the energy functional of Nordström are equivalent with the steady state of a certain nonlinear diffusion-reaction equation. Our digitized formulation of the variational and PDE method in [8] leads to

nonlinear algebraic equations instead of PDE's and the analysis and application of the digital method needs no knowledge on numerical approximation of PDE's.
The new digital filter has the following characteristics.
1. It has a simple fixed filter structure, where the filter coefficients encode the edge information.
2. The digital diffusion-reaction type filter applies to data living on a graph and is therefore useful also for non-flat image features.

The paper is organized as follows. In Section 2 we shortly survey the continuous diffusion-reaction model by Nordström. In Section 3 we introduce our new digital nonlinear filter on a graph adapting some ideas in [4]. The properties and convergence of the new iteration method is considered in Section 4. Finally, we compare the denoising properties of our filter with different image denoising methods in the literature by a numerical example.

§2. The diffusion-reaction filter of Nordström

In this section, we shortly describe the continuous diffusion-reaction filter of Nordström. It can be seen as origin of our digital filter, which will be introduced in the next section.

Let $f : \Omega \to \mathbb{R}$ be the noise contaminated version of a signal $u : \Omega \to \mathbb{R}$, i.e.,

$$f(x) = u(x) + n(x) \qquad x \in \Omega,$$

where n denotes random noise with mean value 0 and variance σ^2,

$$\textstyle\int_\Omega n(x)\,dx = 0, \quad \int_\Omega n(x)^2\,dx = \sigma^2.$$

Usually, Ω denotes a rectangle in $\mathbb{R}^2$, i.e., $\Omega = [0, N_1] \times [0, N_2]$. In order to find an approximation of u, an energy functional minimization problem is considered. Given the noisy f, Nordström suggests to minimize a functional of the form

$$F(\omega_u, u) = \int_\Omega \kappa\Phi(\omega_u(x)) + \beta(u(x) - f(x))^2 + \omega_u(x)\|\nabla u(x)\|^2 dx,$$

where $u : \Omega \to \mathbb{R}$ is the reconstructed image function, and where the control function $\omega_u : \Omega \to (0, 1]$ is supposed to be a fuzzy edge representation which approaches 1 for smooth regions of u and is close to 0 at edges. The purpose of the *edge cost* $\kappa \int_\Omega \Phi(\omega_u(x))dx$ is to impose an explicit penalty of the presence of edges. We suppose that the edge cost coefficient κ is a constant and that the the edge cost density function Φ is twice differentiable. The *deviation cost* $\beta \int_\Omega (u(x) - f(x))^2 dx$ ensures, that u remains to be a faithful approximation of f. Finally the *stabilizing cost* $\int_\Omega \omega_u(x)\|\nabla u(x)\|^2 dx$ is responsible for smoothing of u.

Setting the first variation of F to zero yields the Euler equations

$$\begin{aligned} \beta(u(x) - f(x)) - \nabla \cdot (\omega_u(x) \nabla u(x)) &= 0 \qquad \forall\, x \in \Omega, \\ \kappa\, \Phi'(\omega_u(x)) + \|\nabla u(x)\|^2 &= 0 \qquad \forall\, x \in \Omega, \\ \omega_u(x)\, \tfrac{\partial u(x)}{\partial e_n} &= 0 \qquad \forall\, x \in \partial\Omega, \end{aligned} \tag{1}$$

where $\nabla\cdot$ denotes the divergence operator and $\frac{\partial u}{\partial e_n}$ is the directional derivative in the direction of the outward normal.

We assume that $\Phi'|_{(0,1]} \to \mathbb{R}_-$ is bijective, and that $\Phi'|_{(1,\infty])} \subset \mathbb{R}_+$. Further, let Φ'' be strictly positive on $(0,1)$ and $\Phi''(1) \geq 0$.

If Φ satisfies the above restrictions we obtain the equations

$$\begin{aligned} \beta(f(x) - u(x)) + \nabla \cdot (\omega_u(x) \nabla u(x)) = 0 \quad \forall x \in \Omega, \\ \omega_u(x) = g(\|\nabla u(x)\|) \quad \forall x \in \Omega, \\ \tfrac{\partial u(x)}{\partial e_n} = 0 \qquad \forall\, x \in \partial\Omega, \end{aligned} \tag{2}$$

where the diffusivity function $g : \overline{\mathbb{R}_+} \to (0,1]$ is given by

$$g(s) = (\Phi'|_{(0,1]})^{-1}(-\tfrac{s^2}{\kappa}) \qquad s \geq 0.$$

Hence, g is a strictly positive decreasing differentiable bijection with $g(0) = 1$ and $\lim_{s\to\infty} g(s) = 0$. For example, for $\Phi(s) = s - \ln s$ it takes the form

$$g(s) = (1 + \frac{s^2}{\kappa})^{-1}, \qquad s \geq 0. \tag{3}$$

The diffusivity function in (3) can be easily recognized as the Perona-Malik diffusivity, see [9]. Equation (2) can also be regarded as the steady state equation of

$$\partial_t u = \nabla \cdot (g(\|\nabla u\|) \nabla u) + \beta(f - u).$$

One motivation to consider the Nordström model is to free the user from the difficulty of finding an appropriate stopping time for the Perona- Malik process. Instead, for a given fixed diffusivity g one has to determine the parameter β. Nordström's method may suffer from the same ill-posedness problems as the Perona-Malik equation obtained for $\beta = 0$ (see e.g. [15]). In fact the functional $F(\omega, u)$ is not convex for Perona-Malik diffusivity and may possess numerous local minima.

§3. Digital filters on graphs

We follow the ideas of Chan, Osher and Shen [4] and consider noisy data living on a graph. We want to restore these data.

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, $u : \Omega \to \mathbb{R}$. We can assign a linear order to all nodes of Ω, $\alpha_1 < \alpha_2 < \ldots < \alpha_D$. The value at node α is denoted by u_α. Then u is completely characterized by the vector $\mathbf{u} = (u_1, \ldots, u_D)^T \in \mathbb{R}^D$. 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

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

Obviously, we have $\frac{\partial u}{\partial e}\big|_\alpha = -\frac{\partial u}{\partial e}\big|_\beta$. Further, let

$$|\nabla_e u| = |\nabla_{\alpha,\beta}\, u| := \frac{|u_\alpha - u_\beta|}{\|e_{\alpha,\beta}\|}$$

(see [4] for a different definition).

Fig. 1. Typical partial graphs for rectangular image domains.

Starting with a given noisy signal $\mathbf{u}^0 = (u_\alpha^0)_{\alpha \in \Omega}$, we propose the following iterative filter $H_\lambda : \mathbf{u}^j \to \mathbf{u}^{j+1}$, which is nonlinear and data-dependent. For any node $\alpha \in \Omega$ let

$$u_\alpha^{j+1} = H_\lambda(\mathbf{u}^j) = \frac{1}{1+\lambda}\Big(u_\alpha^j + \lambda u_\alpha^0 + \tau \sum_{\beta \sim \alpha} \frac{g(|\nabla_{\alpha,\beta}\, u^j|)}{\|e_{\alpha,\beta}\|^2}\,(u_\beta^j - u_\alpha^j)\Big) \quad (4)$$

with

$$\tau = \big(\max_{\alpha \in \Omega} \sum_{\beta \sim \alpha} \|e_{\alpha,\beta}\|^{-2}\big)^{-1}.$$

The diffusivity function $g : \mathbb{R}_+ \to (0, 1]$ used in the filter is assumed to be a monotone decreasing function with $g(0) = 1$ and $\lim_{s \to \infty} g(s) = 0$.

Usual candidates for g are the Perona-Malik diffusivity $g(s) = (1+s^2/\kappa)^{-1}$ or the Charbonnier diffusivity $g(s) = (1 + s^2/\kappa)^{-1/2}$. Considering only nodes being directly connected with the node α by an edge in the filter (4), we do not need to determine any boundary conditions. The complete algorithm for the proposed denoising filter is now given as follows.

1. Assign a linear order to all nodes $\alpha_1 < \alpha_2 < \ldots < \alpha_D$ of Ω and initialize $\mathbf{u}^0 = (u^0_\alpha)_{\alpha\in\Omega}$.
2. For $j = 1, 2, \ldots$
   For $k = 1, \ldots, D$ compute $u^{j+1}_{\alpha_k} = H_\lambda(\mathbf{u}^j)$
   end
   end.

The filter H_λ contains a positive parameter λ, called the fitting parameter. The fitting parameter is responsible for balancing smoothing and data fitting. As we will see, this is analogous to the the continuous regularization method.

§4. Properties and Convergence of the Method

For showing the convergence of the above filtering process we first derive a matrix-vector representation of the iteration scheme. Let $\mathbf{u}^j = (u^j_1, \ldots, u^j_D)^T$, $j = 0, 1, 2, \ldots$. Then the iteration (4) can be written in the form

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

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

$$G^j_{\alpha,\beta} := \begin{cases} -\sum_{\gamma\sim\alpha} \frac{g(|\nabla_{\alpha,\gamma} u^j|)}{\|e_{\alpha,\gamma}\|^2} & \text{for } \beta = \alpha, \\ \frac{g(|\nabla_{\alpha,\beta} u^j|)}{\|e_{\alpha,\beta}\|^2} & \text{for } \beta \sim \alpha, \\ 0 & \text{for } \beta \not\sim \alpha. \end{cases} \tag{5}$$

Hence, for the graph on the right hand side of Figure 1, $\mathbf{G}^j$ has at most 9 entries per row. Introducing the matrix

$$\mathbf{A}^j = \mathbf{A}(\mathbf{u}^j) := \mathbf{I} + \tau\,\mathbf{G}^j, \tag{6}$$

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

$$\mathbf{u}^{j+1} = \tfrac{1}{1+\lambda}(\mathbf{A}^j\,\mathbf{u}^j + \lambda\,\mathbf{u}^0). \tag{7}$$

Let us now consider the properties of the iteration matrix $\mathbf{A}^j = \mathbf{A}(\mathbf{u}^j)$ more closely.

Lemma 1. *Let the diffusivity function g satisfy $0 < g(|s|) \le 1$ for $s \in \mathbb{R}$. Then the iteration matrix $\mathbf{A}^j$ given in (6) satisfies the following properties for all $j = 0, 1, 2, \ldots$.*

1. *$\mathbf{A}^j$ is symmetric, i.e., $\mathbf{A}^j = (\mathbf{A}^j)^T$.*
2. *With $\mathbf{1} := (1, 1, \ldots 1)^T \in \mathbb{R}^d$ we have $\mathbf{A}^j\,\mathbf{1} = \mathbf{1}$.*
3. *We have $\mathbf{A}^j \ge \mathbf{0}$, i.e., all entries of $\mathbf{A}^j$ are non-negative. Moreover for the row sum norm and the spectral norm we have $\|\mathbf{A}^j\|_1 = \|\mathbf{A}^j\|_2 = 1$.*

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 $\mathbf{A}^j$. By

$$\tau \sum_{\gamma\sim\alpha} \tfrac{g(|\nabla_{\alpha,\gamma} u^j|)}{\|e_{\alpha,\gamma}\|^2} \le \tau \sum_{\gamma\sim\alpha} \tfrac{1}{\|e_{\alpha,\gamma}\|^2} \le 1$$

for all $\alpha \in \Omega$ it is ensured that the diagonal entries of $\mathbf{A}^j$ are non-negative, and hence $\mathbf{A}^j \ge \mathbf{0}$. 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 2. *Let the diffusivity function g satisfy $0 < g(|s|) \le 1$ for $s \in \mathbb{R}$, and let μ be the average value of $\mathbf{u}^0$, i.e., $\sum_{\alpha\in\Omega} u_\alpha^0 = \mu D$. Then the vectors $\mathbf{u}^j = (u_\alpha^j)_{\alpha\in\Omega}$ obtained by the iteration scheme (4) (resp. (7)) satisfy*

$$\sum_{\alpha\in\Omega} u_\alpha^j = \mu D.$$

Further, we have

$$\min_{\alpha\in\Omega} u_\alpha^0 \le u_\beta^j \le \max_{\alpha\in\Omega} u_\alpha^0 \tag{8}$$

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

$$\|\mathbf{u}^j\|_2 = (\sum_{\alpha\in\Omega} (u_\alpha^j)^2)^{1/2} < \|\mathbf{u}^0\|_2 \quad \forall\, j = 1, 2, \ldots.$$

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

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

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

Since $\mathbf{A}^j \ge \mathbf{0}$ and $\|\mathbf{A}^j\|_\infty = 1$, the application of $\mathbf{A}^j$ to $\mathbf{u}^j$ is a smoothing procedure, where the elements of $\mathbf{v}^{j+1} = \mathbf{A}^j\mathbf{u}^j$ are convex linear combinations of entries in $\mathbf{u}^j$. Hence $\min_{\alpha\in\Omega} u_\alpha^j \le v_\beta^{j+1} \le \max_{\alpha\in\Omega} u_\alpha^j$ for all $\beta \in \Omega$ and the iteration scheme (4) implies

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

The assertion (8) follows now by induction. Finally, by $\|\mathbf{A}^j\|_2 = 1$ we also have

$$\|\mathbf{u}^{j+1}\|_2 \leq \tfrac{1}{1+\lambda}(\|\mathbf{A}^j\mathbf{u}^j\|_2 + \lambda\|\mathbf{u}^0\|_2) \leq \tfrac{1}{1+\lambda}(\|\mathbf{u}^j\|_2 + \lambda\|\mathbf{u}^0\|_2) \leq \|\mathbf{u}^0\|. \quad \square$$

Lemma 3. *Let $\mathbf{A}^j$, $j \in \{0,1,2,\ldots\}$, be the iteration matrix given in (6). Further, assume that the graph (Ω, E) has boundary elements, i.e., not all nodes $\alpha \in \Omega$ are endpoint 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 $(-1+\epsilon, 1-\epsilon)$.*

Proof: Let us consider the matrix $\mathbf{G} = (G_{\alpha,\beta})_{\alpha,\beta=1}^D$ with

$$G_{\alpha,\beta} = \begin{cases} -\sum_{\gamma\sim\alpha} \frac{1}{\|e_{\alpha,\gamma}\|^2} & \beta = \alpha, \\ \frac{1}{\|e_{\alpha,\beta}\|^2} & \beta \sim \alpha, \\ 0 & \text{elsewhere.} \end{cases}$$

Then the eigenvalue 0 of $\mathbf{G}$ is simple, since $\mathbf{G}$ is by definition weakly diagonal dominant and irreducible since (Ω, E) is assumed to be a connected graph. Hence, deleting the last row and the last column of $\mathbf{G}$, we obtain a weakly diagonal dominant $(D-1) \times (D-1)$- matrix, where the condition of strong diagonal dominance holds at least in one row. Thus $\mathbf{G}$ has rank $D-1$ (see e.g [6], p. 356). Further, $\mathbf{G}$ is negative semidefinite, since the Theorem of Geršgorin (see e.g. [6], p. 344) implies that all nonzero eigenvalues of $\mathbf{G}$ lie in the interval $[-2/\tau, 0)$. From the assumption, that Ω possesses boundary elements, it follows that $-2/\tau$ is not an eigenvalue of $\mathbf{G}$ (see [6], p. 344).

Now choose

$$\begin{aligned} d^j_{\min} &:= \min\{G^j_{\alpha,\beta} : G^j_{\alpha,\beta} \neq 0, \alpha \neq \beta\}, \\ d^j_{\max} &:= \max\{G^j_{\alpha,\beta} : G^j_{\alpha,\beta} \neq 0, \alpha \neq \beta\} \leq 1, \end{aligned}$$

i.e, $d^j_{\min}$ is the smallest non-zero off-diagonal entry of $\mathbf{G}^j$ in (5) and $d^j_{\max}$ is the greatest off-diagonal entry of $\mathbf{G}^j$. Observe that there exists a constant $c > 0$ depending on the diffusivity function g, the graph (Ω, E) and the initial vector $\mathbf{u}^0$ but being independent of j, such that $d^j_{\min} > c > 0$ for all $j \geq 0$, since the arguments of g in the entries of $\mathbf{G}^j$ are bounded (see Corollary 2). Let us now consider the matrices

$$\mathbf{A}^j_{\min} := \mathbf{I} + \tau d^j_{\min}\mathbf{G}, \quad \mathbf{A}^j_{\max} := \mathbf{I} + \tau d^j_{\max}\mathbf{G}.$$

We denote the eigenvalues of $\mathbf{A}^j_{\min}$ and $\mathbf{A}^j_{\max}$ in increasing order by $\lambda_k(\mathbf{A}^j_{\min})$ and $\lambda_k(\mathbf{A}^j_{\max})$ for $k = 1,2,\ldots,D$. By $c < d^j_{\min} \leq d^j_{\max} \leq 1$ it follows from the considerations above, that there exists an $\epsilon > 0$ with

$$|\lambda_k(\mathbf{A}^j_{\min})| \leq 1-\epsilon, \quad |\lambda_k(\mathbf{A}^j_{\max})| \leq 1-\epsilon$$

for $k = 1, \ldots, D-1$ and $\lambda_D(\mathbf{A}^j_{\min}) = \lambda_D(\mathbf{A}^j_{\max}) = 1$. Now the matrix $\mathbf{A}^j - \mathbf{A}^j_{\min} = \tau(\mathbf{G}^j - d^j_{\min}\mathbf{G})$ is symmetric and negative semidefinite. This follows again directly from the Theorem of Geršgorin, keeping in mind that all off diagonal entries of $\mathbf{A}^j - \mathbf{A}^j_{\min}$ are nonnegative, and that $(\mathbf{A}^j - \mathbf{A}^j_{\min})\mathbf{1} = \mathbf{0}$. Analogously, we observe that $\mathbf{A}^j - \mathbf{A}^j_{\max} = \tau(\mathbf{G}^j - d^j_{\max}\mathbf{G})$ is symmetric and positive semidefinite. Finally, denoting the eigenvalues of $\mathbf{A}^j$, $\mathbf{A}^j - \mathbf{A}^j_{\min}$ and $\mathbf{A}^j - \mathbf{A}^j_{\max}$ in increasing order and applying the Theorem of Weyl (see e.g. [6], p. 181), we find for $k = 1, \ldots, D-1$

$$-1 + \epsilon \leq \lambda_1(\mathbf{A}^j - \mathbf{A}^j_{\max}) + \lambda_k(\mathbf{A}^j_{\max}) \leq \lambda_k(\mathbf{A}^j)$$

and

$$\lambda_k(\mathbf{A}^j) \leq \lambda_D(\mathbf{A}^j - \mathbf{A}^j_{\min}) + \lambda_k(\mathbf{A}^j_{\min}) \leq 1 - \epsilon. \quad \square$$

We are now ready to prove convergence of the Perona-Malik iteration process (4) for $\lambda = 0$.

Theorem 4. *Let $(\mathbf{u}^j)_{j\geq 0}$ be the sequence of vectors obtained by the iteration process (4) (resp. (7)) applied to a starting vector $\mathbf{u}^0$ for $\lambda = 0$. Then $(\mathbf{u}^j)_{j\geq 0}$ converges to the average value $\mu\,\mathbf{1}$ with $\mu D = \sum_{\alpha\in\Omega} u^0_\alpha$.*

Proof: Let us consider $\mathbf{r}^j = \mathbf{u}^j - \mu\mathbf{1}$. Then we have by Corollary 2

$$\mathbf{1}^T\mathbf{r}^j = \mathbf{1}^T\mathbf{u}^j - \mu\mathbf{1}^T\mathbf{1} = \mu D - \mu D = 0.$$

Hence, it follows

$$\begin{aligned} \mu\mathbf{1} + \mathbf{r}^{j+1} =: \mathbf{u}^{j+1} &= \mathbf{A}^j\,\mathbf{u}^j = \mathbf{A}^j(\mu\mathbf{1} + \mathbf{r}^j) \\ &= \mu\mathbf{A}^j\mathbf{1} + \mathbf{A}^j\mathbf{r}^j = \mu\mathbf{1} + \mathbf{A}^j\mathbf{r}^j, \end{aligned}$$

i.e., $\mathbf{r}^{j+1} = \mathbf{A}^j\mathbf{r}^j$. 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 3 that

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

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

Remark. The convergence of the Perona-Malik process ($\lambda = 0$) in the discrete case has been claimed already by Weickert (see [15], pp. 97). Observe, that in our proof, we do not need that the matrices $\mathbf{A}^j = \mathbf{A}(\mathbf{u}^j)$ have continuous arguments, and that we do not need the positivity of the diagonal entries of $\mathbf{A}^j$. In [9], a similar iteration process has been considered with a special discontinuous diffusivity function g. In this case convergence to a piecewise constant image has been shown.

Further, assuming that the diffusivity function g is Lipschitz continuous we can show

Theorem 5. *Let $(\mathbf{u}^j)_{j\geq 0}$ be the sequence of vectors obtained by the iteration process (4) (resp. (7)) applied to a starting vector $\mathbf{u}^0$. Further, let the Lipschitz condition*

$$\|\mathbf{A}(\mathbf{u}) - \mathbf{A}(\mathbf{v})\|_1 \leq C\|\mathbf{u} - \mathbf{v}\|_1$$

be satisfied for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^D$ with $\|\mathbf{u}\|_2 \leq \|\mathbf{u}^0\|_2$, $\|\mathbf{v}\|_2 \leq \|\mathbf{u}^0\|_2$ and $\mathbf{1}^T\mathbf{u} = \mathbf{1}^T\mathbf{v} = \mathbf{1}^T\mathbf{u}^0$. Then, the sequence $(\mathbf{u}^j)_{j\geq 0}$ converges for all $\lambda > C\|\mathbf{u}^0\|_1$.

Proof: Since the vectors $\mathbf{u}^j$ are bounded (see Corollary 2), there exists at least a converging partial sequence of $(\mathbf{u}^j)_{j\geq 0}$. We show convergence by observing that $(\mathbf{u}^j)_{j\geq 0}$ is a Cauchy sequence in $\|.\|_1$-norm. For $j > 0$ we find

$$\begin{aligned}
\|\mathbf{u}^{j+1} - \mathbf{u}^j\|_1 &= \tfrac{1}{1+\lambda}\|\mathbf{A}^j\mathbf{u}^j - \mathbf{A}^{j-1}\mathbf{u}^{j-1}\|_1 \\
&\leq \tfrac{1}{1+\lambda}(\|\mathbf{A}^j(\mathbf{u}^j - \mathbf{u}^{j-1})\|_1 + \|(\mathbf{A}^j - \mathbf{A}^{j-1})\mathbf{u}^{j-1}\|_1) \\
&\leq \tfrac{1}{1+\lambda}(\|\mathbf{u}^j - \mathbf{u}^{j-1}\|_1 + C\,\|\mathbf{u}^{j-1}\|_1\,\|\mathbf{u}^j - \mathbf{u}^{j-1}\|_1) \\
&= \tfrac{1+C\|\mathbf{u}^0\|_1}{1+\lambda}\|\mathbf{u}^j - \mathbf{u}^{j-1}\|_1,
\end{aligned}$$

where we have used the Lipschitz condition for the iteration matrices. □

Remark. Observe that a Lipschitz condition as in Theorem 5 is satisfied for smooth diffusivity functions g. In particular, it is satisfied for Perona-Malik diffusivity and the Charbonnier diffusivity. The above result on the convergence of the iteration scheme (4) for sufficiently large λ is in accordance with the results of Nordström for another discretization of the Euler equations (2) in [8]. However, the numerical results suggest convergence of his scheme also for small $\lambda > 0$. The same conclusion can be drawn from the numerical results for our scheme. However the theoretical proof for convergence of (4) for small λ remains to be open.

Finally, we show that our digital diffusion-reaction type filter can be seen as digital analogon of the continuous Euler equations in (2).

Theorem 6. *If the filtering sequence $(\mathbf{u}^j)_{j\geq 0}$ given by the scheme (4) converges to a limit vector $\mathbf{u}^* \in \mathbb{R}^D$, then $\mathbf{u}^*$ satisfies*

$$\frac{\tau}{2}\sum_{e\succ\alpha}\frac{\partial}{\partial e}(-g(|\nabla u^*|)\,\frac{\partial u^*}{\partial e}\big|_\alpha + \lambda(u^0_\alpha - u^*_\alpha) = 0 \quad \forall\alpha \in \Omega.$$

Here, again $e \succ \alpha$ means, that α is a node of e.

Proof: Let $\mathbf{u}^* = (u^*_\alpha)_{\alpha\in\Omega}$ be the limit of the iteration process (4), i.e.

$$u^*_\alpha = \tfrac{1}{1+\lambda}\big(u^*_\alpha + \lambda u^0_\alpha + \tau\sum_{\beta\sim\alpha}\tfrac{g(\|\nabla_{\alpha,\beta}u\|)}{\|e_{\alpha,\beta}\|^2}(u^*_\beta - u^*_\alpha)\big).$$

Hence,

$$\tau \sum_{\beta\sim\alpha} \frac{g(|\nabla_{\alpha,\beta}u^*|)}{\|e_{\alpha,\beta}\|^2}(u^*_\beta - u^*_\alpha) + \lambda(u^0_\alpha - u^*_\alpha) = 0.$$

The assertion of the theorem now follows, if we can show that

$$\sum_{\beta\sim\alpha} \frac{g(|\nabla_{\alpha,\beta}u^*|)}{\|e_{\alpha,\beta}\|^2}(u^*_\beta - u^*_\alpha) = \frac{1}{2}\sum_{e\succ\alpha} \frac{\partial}{\partial e}(-g(|\nabla u^*|)\,\frac{\partial u^*}{\partial e}\Big|_\alpha.$$

Suppose that e is the edge linking α and β. Then the definition of the edge derivative implies

$$\begin{aligned}
&\frac{\partial}{\partial e}(-g(|\nabla_e u^*|)\,\frac{\partial u^*}{\partial e}\Big|_\alpha \\
&= \frac{1}{\|e_{\alpha,\beta}\|}\Big[(-g(|\nabla_e u^*|)\frac{\partial u^*}{\partial e})\Big|_\beta - (-g(|\nabla_e u^*|)\frac{\partial u^*}{\partial e})\Big|_\alpha\Big] \\
&= \frac{1}{\|e_{\alpha,\beta}\|^2}\Big[-g(|\nabla_{\alpha,\beta}u^*|)(u^*_\alpha - u^*_\beta) - (-g(|\nabla_{\alpha,\beta}u^*|))(u^*_\beta - u^*_\alpha)\Big] \\
&= \frac{2}{\|e_{\alpha,\beta}\|^2}\,g(|\nabla_{\alpha,\beta}u^*|)(u^*_\beta - u^*_\alpha). \quad \square
\end{aligned}$$

§5. Numerical Results

In this section, we want to show the performance of our method for denoising of images and compare it to other methods. In particular we consider the digital total variation filter by Chan, Osher and Shen [4] and the four-pixel scheme of Welk, Steidl and Weickert [16].

In the test we consider the performance of the proposed methods for the pepper image (see Figure 2(a)). We added zero-mean Gaussian noise, such that the ratio ρ between standard deviation of the image and the noise is one. Here, the signal-to-noise ratio (SNR) is defined by SNR $= 20\log_{10}\frac{\|f-\overline{f}\|_2}{\|n\|_2}$ with f standing for the ideal image with mean $\overline{f}$ and n representing the noise. Thus the SNR of the noisy image is approximately zero.

Figure 2(b) shows the image contaminated with heavy noise and Figure 2(c) a cut through the image, comparing the noisy image with the original in Figure in 2(a). We applied the above mentioned denoising methods, where everytimes we tried to optimize the parameters in order to get an optimal result for each method. The experiments show that our filter works similarly well as existing methods. We present the denoised image and a cut through the image (below) for the three methods to be compared.

Figure 3(a) is obtained using 15 iterations of the digital TV-filter with optimized fitting parameter $\lambda = 12$ (see [4]). Figure 3(b) shows the result of the four-pixel scheme with a time step $\tau = 0.01$ and 18 iterations. Observe, that this scheme is extremely sensitive with respect to the stopping time. For 30 iteration only an SNR of 10.53 is achieved. Figure 3(c) shows the result of our algorithm using 15 iterations of the scheme (4) with Charbonnier diffusivity $g(s) = (1 + 400s^2)^{-1/2}$ and with $\lambda = 0.045$.

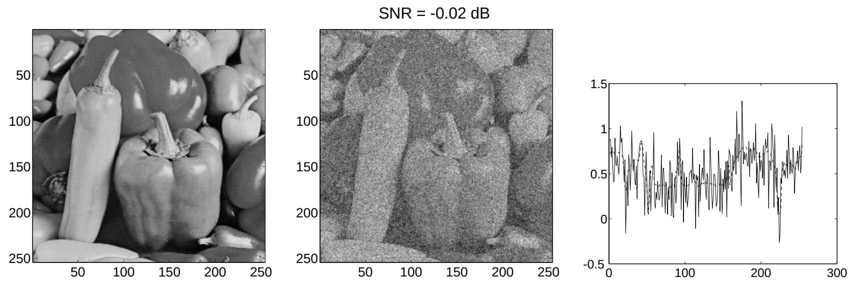

Fig. 2. Original image, noisy image and a cut through the noisy image.

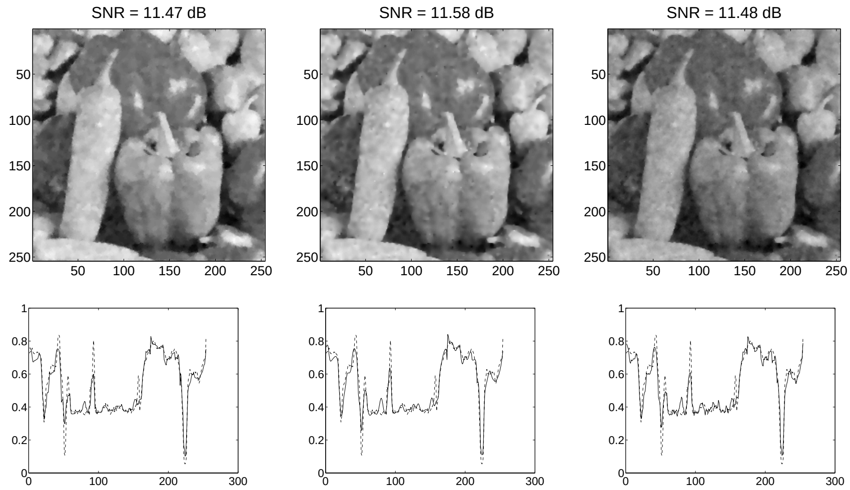

Fig. 3. Denoised images with digital TV filter (left), four-pixel scheme (middle) and digital diffusion-reaction filter (right) and corresponding cuts through the denoised images compared with the original (dashed).

References

1. Black, M. J., G. Sapiro, D. Marimont, and D. Heeger, Robust anisotropic diffusion, IEEE Trans. Image Process. **7** (1998), 421–432.
2. Catté, F., P. L. Lions, J. M. Morel, Image selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal. **29** (1992), 182–193.
3. Chambolle A., R. A. DeVore, N. Lee, and B. L. Lucier, Nonlinear wavelet image processing: variational problems compression and noise removal through wavelet shrinkage, IEEE Trans. Image Process. **7** (1998), 319–335.
4. Chan, T. F., S. Osher, and J. Shen, The digital TV filter and nonlinear denoising, IEEE Trans. Image Process. **10** (2001), 231–241.

5. Charbonnier, P., L. Blanc-Féaud, G. Aubert, and M. Barlaud, Two deterministic half-quadratic regularization algorithms for computed imaging, in *Proc. 1994 IEEE International Conference on Image Processing*, volume 2, IEEE Computer Society Press, Austin, 1994, 168–172.
6. Horn, R. A., and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
7. Ma, J., and G. Plonka, Combined curvelet shrinkage and nonlinear anisotropic diffusion, IEEE Trans. Image Process., to appear.
8. Nordström, N., Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection, Image Vision Comput. **8** (1990), 318–327.
9. Perona, P., and J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Machine Intell. **12** (1990), 629–639.
10. Plonka, G., and G. Steidl, A multiscale wavelet-inspired scheme for nonlinear diffusion, International Journal of Wavelets, Multiresolution and Information Processing **4** (2006), 1–22.
11. Plonka, G., and J. Ma, Convergence of an iterative nonlinear scheme for denoising of piecewise constant images, International Journal of Wavelets, Multiresolution and Information Processing, to appear.
12. Rudin, L., S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D **60** (1992), 259–268.
13. Scherzer, O. and J. Weickert, Relations between regularization and diffusion filtering, Journal of Mathematical Imaging and Vision **12** (2000), 43–63.
14. Steidl, G., 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 J. Numer. Anal., **42** (2004), 686–713.
15. Weickert, J., *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.
16. M. Welk, G. Steidl, and J. Weickert: A four–pixel scheme for singular differential equations. in *Scale-Space and PDE Methods in Computer Vision.*, R. Kimmel, N. Sochen, J. Weickert (eds.), Lecture Notes in Computer Science, Springer, Berlin, 2005, 610–621.

Gerlind Plonka
University of Duisburg-Essen, Department of Mathematics
47048 Duisburg, Germany
`gerlind.plonka@uni-due.de`
`http://www.uni-duisburg.de/FB11/STAFF/PLONKA/home.html`