International Journal of Wavelets, Multiresolution and Information Processing

CONVERGENCE OF AN ITERATIVE NONLINEAR SCHEME FOR DENOISING OF PIECEWISE CONSTANT IMAGES

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

JIANWEI MA
Laboratoire LMC-IMAG, University Joseph Fourier, BP 53,
38041 Grenoble Cedex 9, France
jianwei.ma@imag.fr

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

In this paper we present a new efficient iterative nonlinear scheme for recovering of a piecewise constant image from an observed image containing additive noise.
We apply an adaptive neighborhood filter which comes from robust statistics and completely rejects outliers being greater than a certain constant. We prove that the iterated application of the scheme leads to a piecewise constant image. This observation generalizes the known results on convergence of nonlinear diffusion schemes to a constant steady-state. Moreover, we show that the partition of the image determining the piecewise constant steady-state after an infinite iteration process can already be found after a finite number of iteration steps. This result can be used for a fast approximation of the piecewise constant image by a mean value procedure. We examine the relations of our scheme to average and bilateral filtering, diffusion filtering and wavelet shrinkage. Numerical experiments illustrate the performance of the algorithm.

Keywords: Nonlinear diffusion; bilateral filter; discontinuity-preserving; denoising; wavelets.

AMS Subject Classification: 65T60, 65M06, 65M12, 94A12

1. Introduction

Pre-smoothing and noise removal are important tools of image preprocessing in order to improve the performances of image compression, detection enhancement etc.. In a wide variety of applications, the images are discontinuous, and the challenge is to smooth an image while preserving its edges. There are different methods in the literature to tackle this problem where mostly local and adaptive schemes are applied. Such methods can be based on anisotropic diffusion (see e.g. [6, 23, 26, 33, 34, 35]), robust statistics (see e.g. [3, 17, 16, 27, 4, 5]), averaging and bilateral

filtering (see e.g. [32, 2, 11, 15]), regularization techniques (see e.g. [24, 1, 9]), tools from Harmonical Analysis (see e.g. [8, 12, 30]) and others.

Close connections between these methods have been examined (see e.g. [2, 15, 3, 7, 22, 29, 31]).

Typical averaging local filters like those based on explicit discrete anisotropic diffusion use the data of the image to be processed only as input in the first iteration while in all further iteration steps the original data are not involved. Each step of filtering gradually removes noise and details from the data, and one crucial question is to find the right stopping time for the filtering in order to obtain the optimal 'restoration' result (see e.g. [23, 33, 35, 34]). Most of these iterative schemes are known to converge to a constant steady-state, i.e., to a constant image which is the spatial average of the original intensities (see e.g. [33]).

In contrast to these methods, the iterative regularization algorithms consider a data term and a smoothing term in each iteration step. Similar models are also used in the theory of W-smoothers coming from robust statistics. A very well working example of a discrete regularization method is the digital TV filter of Chan, Osher and Shen [9].

Another statistical iterative method by Polzehl and Spokoiny [27] is a smoothing scheme using weights which are computed with the help of adaptive estimates of the local variances of the noise during the iteration. Further, in a control step the obtained results are compared with smoothed images found in former iteration steps in order to move not too far from the initial image.

Some methods do not work iteratively at all but apply a global smoothing filter as e.g. the method of nonlocal means by Buades, Coll and Morel [4, 5] or the original bilateral filter proposed by Tomasi and Manduchi [32].

Close relations between image smoothing methods can be especially observed by comparison of the related numerical schemes for the digital signals and images. In the last time, a lot of hybrid algorithms have been proposed in order to improve the performance of image denoising (see e.g. [14, 13, 25, 19, 35]).

In this paper we present a simple and efficient nonlinear algorithm for recovering of a piecewise constant image from an observed image containing additive noise. The complete denoising scheme consists of three steps.

In a first step we apply an adaptive neighborhood filtering scheme. The proposed nonlinear averaging filter is a typical smoothing filter which has no recourse to the input data during the iteration process. In contrast to the usually taken diffusion filters (like Perona-Malik-filter, Charbonnier-filter, regularized TV-filter etc.) the filter in our scheme can be seen as a so-called "robust" filter, i.e., it completely rejects outliers being greater than a certain constant. The used filter corresponds to the so-called cup-function used by Winkler et al. [36], which is a robust prior function (see also [21, 17]). Other robust filters with this property are e.g. the Tukey's biweight used for robust anisotropic diffusion [3] and the truncated modulus [36].

We will be able to show that the iterated application of our proposed neighborhood filter leads to a piecewise constant image, i.e., there exists a partition of the image, such that in each subdomain of the partition the spatial average of the pixel values of all pixels belonging to this subset is found. This observation generalizes the known result on convergence of the image to the constant steady-state mentioned above. We will show even more, namely that the partition of the image determining the piecewise constant steady-state after an infinite iteration process can already be found after a finite number of iteration steps. Each iteration step of the proposed scheme can be seen as a matrix vector multiplication where the iteration matrix is sparse and can be simply computed. These iteration matrices determine already certain partitions of the image into subdomains. Hence we only need to check whether the partition changes further after one iteration step by comparing the iteration matrices and stop the process if the difference between two such matrices vanishes or is small enough.

In the second step we apply a simple mean value procedure, which maps the mean value of intensities in a subdomain of the partition to each pixel of this subdomain. If the final partition has been obtained already in the iteration procedure then this second step yields the same result as an application of infinitely many iteration steps of the filtering scheme.

If the signal to noise ratio is rather low, then we apply an outlier removement procedure in the third step, where all subdomains of the partition with less than a certain number of pixels are regarded as outliers and we compute local median values for these pixel values.

The paper is organized as follows. In Section 2 we introduce the iterative neighborhood scheme and investigate its properties and its convergence.

In Section 3 the relations of the iterative scheme to other methods of image denoising are considered. We examine the relations of this new filter to average and bilateral filtering, diffusion filtering and wavelet shrinkage. We observe that our scheme can be interpreted as an iterative bilateral filter with a special weight function. It can also be seen as a discretization of a special diffusion equation with a diffusivity function which has been not used so far in this context, and as an application of Haar wavelet frames with hard shrinkage.

In Section 4 we describe the complete algorithm and discuss the choice of parameters.

The numerical results impressively show the performance of the algorithm. We compare the proposed algorithm with some other image denoising methods considered in the literature. In particular, we also consider the digital TV filter by Chan, Osher and Shen [9] as an example of a regularization method, the four-pixel scheme by Welk, Steidl and Weickert [35] as one example of TV-diffusion which can be also interpreted as a Haar wavelet shrinkage scheme, the adaptive weights smoothing by Polzehl and Spokoiny [27] and the method of nonlocal means by Buades, Coll and Morel [4, 5]. Observe, that our algorithm is especially suitable for recovering of piecewise constant images. It works very efficient for images containing thin

structures, since it is less smoothing than other methods in the neighborhood of edges.

2. The iterative algorithm and its convergence

We consider a piecewise constant digital image $f = (f_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$. The aim is to reconstruct f from an observed image $u^0 = (u^0_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$, where

$$u^0 = f + n,$$

with $n = (n_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$ being a vector of normally distributed noise $n_{i,j} \in N(0, \sigma^2)$ with vanishing mean value and variance σ^2. Usually the variance σ^2 of the noise is unknown.

For the reconstruction of f we propose the following simple neighborhood filter iteration scheme as a first step.

Iteration scheme

Put a suitable shrinkage parameter $\theta > 0$, a smoothing parameter $0 < \alpha \leq \frac{1}{6}$, and let $u^0 = (u^0_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$ be the observed image.

Perform the following iteration for $k = 0, 1, \ldots$.

For $i = 0, \ldots, N_1 - 1$ and $j = 0, \ldots, N_2 - 1$ compute the new image values

$$u^{k+1}_{i,j} = u^k_{i,j} + \alpha \sum_{\substack{(l,m)\in N(i,j)\\(l,m)\neq(i,j)}} \frac{T_\theta(u^k_{l,m} - u^k_{i,j})}{(l-i)^2 + (m-j)^2}, \tag{2.1}$$

where $T_\theta(x)$ denotes the function

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

Further, $N(i,j)$ denotes a neighborhood of the pixel (i,j), where we use periodic boundary conditions.

The parameter α controls the smoothing process in each iteration. Taking e.g.

$$N(i,j) := \{(l,m) : \ |i-l| \leq 1; \ |j-m| \leq 1\},$$

we obtain

$$u^{k+1}_{i,j} = u^k_{i,j} + \alpha \sum_{\substack{r,s=-1\\(r,s)\neq(0,0)}}^{1} \frac{T_\theta(u^k_{i+r,j+s} - u^k_{i,j})}{r^2 + s^2}. \tag{2.2}$$

The coefficients of this scheme are illustrated in Figure 1. Other schemes with greater neighborhood can be simply derived.

The scheme proposed above can be related to ideas of adaptive smoothing (see [2, 28]), bilinear filtering (see [2, 32]) anisotropic diffusion filtering (see [33, 34, 35])

$$\begin{array}{ccc} \frac{\alpha}{2} & \alpha & \frac{\alpha}{2} \\ \alpha & 1-6\alpha & \alpha \\ \frac{\alpha}{2} & \alpha & \frac{\alpha}{2} \end{array} \qquad\qquad \begin{array}{ccc} 0 & 0 & \frac{\alpha}{2} \\ 0 & 1-3\alpha & \alpha \\ \frac{\alpha}{2} & \alpha & 0 \end{array}$$

Fig. 1. Illustration of the coefficients in the scheme (2.2) for the case that the difference in gray value between neighboring pixels are all smaller than θ (left) and for the case that the difference in gray value is greater than θ for some neighbors (right).

as well as iterative Haar wavelet shrinkage (see [22, 25]). We consider these relations in detail in Section 3.

Let us now consider the properties and convergence of the iteration scheme. For convenience we restrict our considerations to the scheme (2.2). All ideas can be simply transferred to more complex schemes.

We choose one index for pixel numbering of the digital image u^k. Put $N := N_1 \cdot N_2$ and

$$n = i + N_1 j, \qquad i = 0, \dots, N_1 - 1, \quad j = 0, \dots, N_2 - 1,$$

such that the pixel n corresponds to (i, j). Then the iteration scheme (2.2) can be written in matrix-vector form as

$$u^{k+1} = \mathbf{A}^k u^k,$$

where $u^k = (u_0^k, \dots, u_{N-1}^k)^T$ and where $\mathbf{A}^k = (A_{n,p}^k)_{n,p=0}^{N-1} \in \mathbb{R}^{N \times N}$ is a sparse matrix of the form

$$A_{n,p}^k := \tag{2.3}$$

$$\begin{cases} 1 - \kappa_n \alpha & \text{for } p = n, \\ \alpha & \text{for } p \in \{n-1,\, n+1,\, n-N_1,\, n+N_1\} (\mathrm{mod}\ N) \text{ and } |u_n^k - u_p^k| < \theta, \\ \alpha/2 & \text{for } p \in \{n-1+N_1,\, n+1+N_1,\, n+1-N_1,\, n-1-N_1\} (\mathrm{mod}\ N) \\ & \text{and } |u_n^k - u_p^k| < \theta, \\ 0 & \text{elsewhere.} \end{cases}$$

Here κ_n (with $0 \le \kappa_n \le 6$) is chosen such that the sum of entries in the nth row of $\mathbf{A}^k$ is 1.

Now we observe the following properties of the iteration matrix $\mathbf{A}^k$.

1. The number of nonzero entries in each row (column) of $\mathbf{A}^k$ is at most 9.
2. For $\alpha \le 1/6$ all entries of $\mathbf{A}^k$ are non-negative, i.e., $\mathbf{A}^k \ge \mathbf{0}$.
3. With $\mathbf{1} := (1, \dots, 1)^T \in \mathbb{R}^N$ we have $\mathbf{A}^k \mathbf{1} = \mathbf{1}$.
4. The iteration matrix $\mathbf{A}^k$ is symmetric, i.e. $\mathbf{A}^k = (\mathbf{A}^k)^T$.

5. For $\alpha \leq 1/6$ the Properties 2 and 3 imply for the row sum norm $\|\mathbf{A}^k\|_\infty = 1$ and in particular $\rho(\mathbf{A}^k) = 1$ where $\rho(\mathbf{A}^k)$ denotes the spectral radius of $\mathbf{A}^k$.
6. By suitable ordering of rows and columns of $\mathbf{A}^k$ we can transfer $\mathbf{A}^k$ into a block diagonal matrix, where the nth and the pth row (column) belong to the same block if there exists a sequence of indices n_1, $n_2, \ldots, n_\nu$ such that $n_1 \in N(n)$, $n_2 \in N(n_1), \ldots, p \in N(n_\nu)$ and if

$$|u_n^k - u_{n_1}^k| < \theta,\ |u_{n_1}^k - u_{n_2}^k| < \theta, \ldots |u_{n_\nu}^k - u_p^k| < \theta$$

holds. Let J_k be the number of such blocks. We assume that this ordering is done in the same way for rows and columns, i.e., there is a permutation matrix $\mathbf{P}^k \in \mathbb{R}^{N\times N}$ such that

$$\text{block}\,(\mathbf{A}_1^k, \ldots, \mathbf{A}_{J_k}^k) = (\mathbf{P}^k)^T\, \mathbf{A}^k\, \mathbf{P}^k.$$

Then for each matrix block $\mathbf{A}_m^k$, $m = 1, \ldots, J_k$, the properties 1 – 5 are again satisfied. Moreover, each of the block matrices $\mathbf{A}_m^k$ is irreducible by construction.

The ordering of $\mathbf{A}^k$ into blocks defines a partition S^k of the pixel set $\{0, \ldots, N-1\}$ into different sets, $S^k = \{S_m^k\}_{m=1}^{J_k}$ separating pixels whose pixel values differ strongly enough. Two pixels n, n' are in the same set S_m^k if in the above procedure the nth and the n'th row of $\mathbf{A}^k$ are transferred into the same block matrix $\mathbf{A}_m^k$.
7. Finally we observe:

Lemma 2.1. *For $\alpha \leq 1/6$ each block matrix $\mathbf{A}_m^k$, $m = 1, \ldots, J_k$ of the iteration matrix $\mathbf{A}^k$ in (2.3) possesses only one simple eigenvalue 1 and no further eigenvalues on the unit circle.*

Proof. 1. Consider for $m \in \{1, \ldots, J_k\}$ the matrix $\mathbf{B}_m^k = (B_{\mu,\nu}^k) = \mathbf{A}_m^k - \mathbf{I}$, where $\mathbf{I}$ denotes the unit matrix of suitable size. Denote the size of $\mathbf{B}_m^k$ by $N_{k,m}$. Then, by $\mathbf{A}_m^k \geq \mathbf{0}$ and $\mathbf{A}_m^k\, \mathbf{1}_m = \mathbf{1}_m$ (with $\mathbf{1}_m$ the 1-vector of length $N_{k,m}$) we have $B_{\mu,\mu} \leq 0\ \forall \mu$ and $B_{\mu,\nu} \geq 0\ \forall\, \mu, \nu$ with $\mu \neq \nu$ as well as $\mathbf{B}_m^k \mathbf{1}_m = \mathbf{0}$. Hence, $\mathbf{B}_m^k$ is weakly diagonal dominant. Further, in all main submatrices $(B_{\mu,\nu}^k)_{\mu,\nu=0}^j$, $j = 0, \ldots, N_{k,m} - 2$ (up to the full matrix $\mathbf{B}_m^k$) there is at least one row, where the condition of strong diagonal dominance holds since the block matrix $\mathbf{A}_m^k$ (and hence $\mathbf{B}_m^k$) is irreducible by definition. Thus, all main submatrices of $\mathbf{B}_m^k$ (up to the full matrix) are invertible (see e.g. [18], p. 363), i.e., $\text{rk}\,(\mathbf{B}_m^k) = N_{k,m} - 1$. We conclude that the eigenvalue 1 of $\mathbf{A}_m^k$ is simple.

2. Since $\mathbf{A}_m^k$ is symmetric and $\rho(\mathbf{A}_m^k) = 1$, we only need to check, whether $\mathbf{A}_m^k$ has the eigenvalue -1. By construction, $\mathbf{A}_m^k + \mathbf{I}$ is strongly diagonal dominant and hence invertible such that -1 can not be an eigenvalue of $\mathbf{A}_m^k$. □

As we have seen in Property 6, each iteration matrix $\mathbf{A}^k$ uniquely defines a partition $S^k = \{S_m^k\}_{m=1}^{J_k}$ of the pixel set. Let r_k be the greatest distance between

two pixel values in the same set S_m^k, $m \in \{1, \ldots, J_k\}$, i.e.,

$$r_k := \max\{|u_n^k - u_{n'}^k|;\ n,\, n' \in S_m^k,\, m \in \{1, \ldots, J_k\}\},$$

and let

$$\mu_m^k = \frac{1}{\#S_m^k} \sum_{n \in S_m^k} u_n^k, \qquad m = 1, \ldots, J_k \tag{2.4}$$

be the mean value of pixel values in S_m^k. Here $\#S_m^k$ denotes the number of elements in the set S_m^k.

Now let us consider the question, how the partition S^k can change after one iteration step of the iteration scheme, i.e., how the partitions S^k and S^{k+1} are related?

If the new partition S^{k+1} is different from S^k, then for a fixed set $S_m^k \in S^k$ the following types of change are possible:
1. S_m^k is divided into two or more subsets $S_{m_1}^{k+1}, S_{m_2}^{k+1}, \ldots$.
2. S_m^k is united with one or more neighboring sets, i.e., $S_{m_1}^{k+1} = S_m^k \cup S_{m'}^k \cup \ldots \cup S_{m^\nu}^k$.
3. S_m^k is united with one or more neighboring subsets of $S_{m'}^k$ to obtain a new set $S_{m_1}^{k+1}$ of S^{k+1}.
4. S_m^k is divided into two or more subsets $S_{m_1}^{k+1}, \ldots$ which are united themselves with one ore more neighboring sets $S_{m'}^k$ (or a neighboring subsets of $S_{m'}^k$).

Example 2.1. We give two small examples for the occurrence of change of partitions S^k of type 1 and type 2.

(i) Consider a (3×3) image u^0 of the form

$$u^0 = \begin{pmatrix} 0 & 0 & 2\theta - 2\epsilon \\ 0 & \theta - \epsilon & 2\theta - 2\epsilon \\ 0 & 0 & 2\theta - 2\epsilon \end{pmatrix}$$

with θ being a fixed shrinkage parameter and with $0 < \epsilon < \theta$. Let the index numbering start at the left upper corner of the matrix (image), i.e., $u_0^0 = u_1^0 = 0$, $u_2^0 = 2\theta - 2\epsilon$, $u_3^0 = 0$, etc.. The corresponding partition $S^0 = \{\{0, 1, 2, 3, 4, 5, 6, 7, 8\}\}$ consists of only one set since the fourth pixel value $\theta - \epsilon$ is obviously connected with all eight neighbors. Application of one iteration of the scheme (2.2) yields

$$u^1 = \begin{pmatrix} \frac{\alpha}{2}(\theta - \epsilon) & \alpha(\theta - \epsilon) & (2 - \frac{\alpha}{2})(\theta - \epsilon) \\ \alpha(\theta - \epsilon) & (1 - 2\alpha)(\theta - \epsilon) & (2 - \alpha)(\theta - \epsilon) \\ \frac{\alpha}{2}(\theta - \epsilon) & \alpha(\theta - \epsilon) & (2 - \frac{\alpha}{2})(\theta - \epsilon) \end{pmatrix}.$$

Since $\alpha \leq \frac{1}{6}$, it follows for $\epsilon < \frac{\alpha}{1+\alpha}\theta$ that $|((1 - 2\alpha) - (2 - \alpha))(\theta - \epsilon)| > \theta$. Hence the corresponding new partition $S^1 = \{S_1^1,\, S_2^1\}$ with $S_1^1 = \{0, 1, 3, 4, 6, 7\}$, $S_2^1 = \{2, 5, 8\}$ consists of two sets.

(ii) Let again θ be a fixed shrinkage parameter and let $0 < \epsilon < \theta$. Consider now a (4×4) image u^0 of the form

$$u^0 = \begin{pmatrix} 0 & \theta+\epsilon & \theta+\epsilon & \theta+\epsilon \\ 0 & \theta+\epsilon & 2\epsilon & \theta+\epsilon \\ 0 & \theta+\epsilon & \theta+\epsilon & \theta+\epsilon \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

yielding the partition $S^0 = \{\{0,4,8,12,13,14,15\},\{1,2,3,5,6,7,9,10,11\}\}$ consisting of two sets. Application of one iteration of the scheme (2.2) leads to

$$u^1 = \begin{pmatrix} 0 & \theta(1-\frac{\alpha}{2})+\epsilon(1+\frac{\alpha}{2}) & \theta(1-\alpha)+\epsilon(1+\alpha) & \theta(1-\frac{\alpha}{2})+\epsilon(1+\frac{\alpha}{2}) \\ 0 & \theta(1-\alpha)+\epsilon(1+\alpha) & 6\alpha\theta+\epsilon(2-6\alpha) & \theta(1-\alpha)+\epsilon(1+\alpha) \\ 0 & \theta(1-\frac{\alpha}{2})+\epsilon(1+\frac{\alpha}{2}) & \theta(1-\alpha)+\epsilon(1+\alpha) & \theta(1-\frac{\alpha}{2})+\epsilon(1+\frac{\alpha}{2}) \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now for $\epsilon < \frac{\alpha}{1+\alpha}\theta$, the new partition only consists of one set $S^1 = \{\{0,1,2,3,4,5,6,7,8,9,10,11,\ 12,13,14,15\}\}$. The above two examples also nicely show the influence of the parameter α. Choosing α as big as possible encourages the change of partitions in the above two configurations.

After these considerations we prove the following convergence theorem which is the main result of this paper.

Theorem 2.1. *Consider the iteration scheme (2.2) with $\alpha \le 1/6$ and starting with the noisy image u^0. Then, after a finite number of k iterations the partition $S^k = (S^k_m)_{m=1}^{J_k}$ of pixels corresponding to the blocks $\mathbf{A}^k_m$, $m = 1,\dots,J_k$ of the ordered iteration matrix $\mathbf{A}^k$ is settled and will not change further, i.e., there exists a finite k such that*

$$S^k = S^{k+\nu} \quad \nu = 1,2,\dots.$$

The iteration scheme (2.2) will converge to the spatial average of u^k in all subsets S^k_m,

$$\lim_{\nu\to\infty} u^{k+\nu} = \sum_{m=1}^{J_k} \mu^k_m \, \chi_{S^k_m}$$

with μ^k_m in (2.4) and where $\chi_{S^k_m}$ is the characteristic function corresponding to the index set S^k_m, i.e.,

$$\chi_{S^k_m}(n) = \begin{cases} 1 \text{ for } n \in S^k_m, \\ 0 \text{ for } n \notin S^k_m. \end{cases}$$

Proof. Let u^k be the vector obtained after the kth iteration of (2.2) and let $\mathbf{A}^k$ be the next iteration matrix. Further, let $S^k = (S^k_m)_{m=1}^{J_k}$ be the partition of the index set corresponding to

$$\text{block}\,(\mathbf{A}^k_1,\dots,\mathbf{A}^k_{J_k}) = (\mathbf{P}^k)^T\,\mathbf{A}^k\,\mathbf{P}^k$$

as described above, where each of the matrices $\mathbf{A}_m^k$, $m = 1, \ldots, J_k$ is symmetric, non-negative and has a simple eigenvalue 1 with $\mathbf{A}_m^k \, \mathbf{1}_m = \mathbf{1}_m$ with $\mathbf{1}_m = (1, \ldots, 1)^T$ being a vector of suitable length. Moreover, the above considerations imply that there is an ϵ with $0 < \epsilon \leq 1$ such that all further eigenvalues λ_j of $\mathbf{A}^k$ are bounded by $|\lambda_j| < 1 - \epsilon$.

By $\|\mathbf{A}^k\|_2 = 1$ we obviously have $\|u^{k+1}\|_2 = \|\mathbf{A}^k u^k\|_2 \leq \|u^k\|_2$, where $\|\cdot\|_2$ denotes the spectral norm for matrices and the Euclidean norm for vectors.

Let now $\{e_1^k, \ldots, e_{J_k}^k\}$ be the set of eigenvectors of $\mathbf{A}^k$ corresponding to the J_k eigenvalues 1, being obtained by extending the eigenvectors $\mathbf{1}_m$ of $\mathbf{A}_m^k$, $m = 1, \ldots, J_k$ to vectors of length N by inserting zeros. Now, we extend this orthogonal set $\{e_1^k, \ldots, e_{J_k}^k\}$ to an orthogonal set of eigenvectors $\{e_1^k, \ldots, e_{J_k}^k, b_1^k, \ldots, b_{N-J_k}^k\}$ of $\mathbf{A}^k$, where again the vectors b_j^k are constructed by extending the eigenvectors of all block matrices $\mathbf{A}_m^k$ to eigenvectors of $\mathbf{A}^k$ by inserting zeros.

This set of eigenvectors now forms an orthogonal basis of $\mathbb{R}^N$. Consider the unique decomposition of u^k with respect to this basis,

$$u^k = \sum_{m=1}^{J_k} \mu_m^k \, e_m^k + \sum_{l=1}^{N-J_k} c_l \, b_l^k = e^k + r^k, \quad c_l \in \mathbb{R},$$

where e^k and r^k represent the first and the second sum, respectively. Hence,

$$\|u^k\|_2^2 = \|e^k\|_2^2 + \|r^k\|_2^2.$$

We observe that

$$u^{k+1} = \mathbf{A}^k \, u^k = \sum_{m=1}^{J_k} \mu_m^k \, e_m^k + \sum_{l=1}^{N-J_k} c_l \, \mathbf{A}^k \, b_l^k = e^k + r^{k+1}$$

and with $\mathbf{A}^k b_l^k = \lambda_l \, b_l^k$ we obtain

$$\|r^{k+1}\|_2^2 = \sum_{l=1}^{N-J_k} |c_l|^2 |\lambda_l|^2 \, \|b_l^k\|_2^2 \leq (1-\epsilon)^2 \|r^k\|_2^2. \tag{2.5}$$

Moreover, in each subset S_m^k we have

$$\|r_m^{k+1}\|_2 \leq (1-\epsilon) \|r_m^k\|_2, \tag{2.6}$$

where the vector r_m^k is obtained by removing all components of the vector r^k not belonging to the index set S_m^k.

Observe that $\|u^k\|_2^2$ is bounded from below by $\mu \, \|\mathbf{1}\|_2^2 = \mu \, N$, where $\mu = \frac{1}{N} \sum_{n=0}^{N-1} u_n^k = \frac{1}{N} \sum_{n=0}^{N-1} u_n^0$. Further, the above considerations imply that

$$\|u^{k+1}\|_2^2 \leq (1-\epsilon) \, \|r^k\|_2^2 + \|e^k\|_2^2 = \|u^k\|_2^2 - \epsilon \, \|r^k\|_2^2.$$

By (2.5), which holds for each iteration independently of the special form of the iteration matrix, there exists an integer k_1 such that $\|r^{k_1}\|_2 < 6\alpha\theta$. Let J_{k_1} be the number of index sets after k_1 iterations. Since the application of the scheme (2.2) ensures that $|u_n^k - u_n^{k+1}| < 6\alpha\theta$, we simply see, that a possible change of index sets

for a k-th iteration with $k \geq k_1$ can be only of type 2. Each change of index sets of type 2 decreases the number J_k of different index sets by at least one. Since we have started with a finite number of sets, there will be no further change of index sets of type 2 after a finite number of iterations. Now, from (2.6) the further assertion of the theorem follows. □

Remark. As we can see in the proof of Theorem 2.1, only the properties 2-7 of the iteration matrices $\mathbf{A}^k$ are applied for this convergence result. These properties can also be derived if the weights in the iteration scheme (2.2) are taken in another way, as long as the neighborhood filter is local and the radiometric weight punishes big intensity differences by having no influence to the smoothing process. For example, the function T_θ can be also replaced by Tukey's biweight

$$T_\sigma(x) = \begin{cases} x(1-\frac{x^2}{\sigma^2})^2 & |x| \leq \sigma, \\ 0 & |x| > \sigma \end{cases}$$

with a suitable fixed parameter σ.

3. Relations to other schemes

In this section we shortly describe the close relations of our neighborhood filter introduced in Section 2 to other schemes in the literature.

3.1. *Adaptive smoothing and bilateral filtering*

Given an image $u^k = (u^k_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$, one iteration of *adaptive smoothing* as described in [28] yields

$$u^{k+1}_{i,j} = \frac{\sum_{r=-1}^{1}\sum_{s=-1}^{1} u^k_{i+r,j+s}\,\omega^k_{i+r,j+s}}{\sum_{r=-1}^{1}\sum_{s=-1}^{1}\omega^k_{i+r,j+s}},$$

with the weight

$$w^k_{i+r,j+s} = \exp\left(\frac{-|d^k_{i,j}|}{2\sigma^2}\right), \tag{3.1}$$

where σ is a suitable constant, and where $d^k_{i,j}$ depends on the magnitude of the gradient of $u^k_{i,j}$, e.g.

$$d^k_{i,j} = \tfrac{1}{\sqrt{2}}\left((u^k_{i+1,j}-u^k_{i,j})^2 + (u^k_{i,j+1}-u^k_{i,j})^2 + (u^k_{i-1,j}-u^k_{i,j})^2 + (u^k_{i,j-1}-u^k_{i,j})^2\right)^{\frac{1}{2}},$$

cf. [28].

Bilateral filtering was introduced in [32] as a nonlinear filter which combines domain and range filtering. The discrete version of bilateral filtering can be written as follows. For a given image $u^0 = (u^0_{i,j})_{i=0,j=0}^{N_1-1,N_2-1}$ one computes u^1 by

$$u^1_{i,j} = \frac{\sum_{r=-n}^{n}\sum_{s=-n}^{n} u^0_{i+r,j+s}\,\omega_{i+r,j+s}}{\sum_{r=-n}^{n}\sum_{s=-n}^{n}\omega_{i+r,j+s}}, \tag{3.2}$$

where the weight $\omega_{i+r,j+s}$ is given as a product of a spatial and a radiometric weight,

$$\omega_{i+r,j+s} = (\omega_1)_{i+r,j+s} \cdot (\omega_2)_{i+r,j+s}.$$

The first weight measures the geometric distance between the center sample (i,j) and the sample $(i+r,j+s)$ using the Euclidean metric. The second weight measures a radiometric distance between the center sample $u^0_{i,j}$ and $u^0_{i+r,j+s}$. In [32], it is suggested to apply only one iteration of type (3.2) with

$$(\omega_1)_{i+r,j+s} = \exp\left(-\frac{r^2+s^2}{2\sigma_1^2}\right), \qquad (\omega_2)_{i+r,j+s} = \exp\left(-\frac{(u^0_{i+r,j+s}-u^0_{i,j})^2}{2\sigma_2^2}\right).$$

The bilateral filtering considered in [32] is not an iterative procedure. We remark that iterative bilateral filtering can be seen as a generalization of the adaptive smoothing algorithm (see e.g. [28, 2]). Further improvements for image denoising have been achieved using these ideas (see e.g. [11]).

Our scheme (2.1) is related to the two schemes (3.1) and (3.2) as follows. We also take a weighted average of a pixel value and its neighbors in (2.2). In our approach, the weight depends on the size $|u^k_{i+r,j+s} - u^k_{i,j}|$ as well as on the distance of pixels $(i+r,j+s)$ and (i,j) analogously as in (3.2) and in contrast to the weight in (3.1).

In fact, our scheme (2.1) can be seen as an iterative bilateral filtering with the weight functions

$$(\omega_1)^k_{i+r,j+s} = \begin{cases} \frac{\alpha}{r^2+s^2} & (r,s) \neq (0,0), \\ 1 & (r,s) = (0,0), \end{cases}$$

$$(\omega_2)^k_{i+r,j+s} = \begin{cases} 1 & |u^k_{i+r,j+s} - u^k_{i,j}| < \theta \text{ and } (r,s) \neq (0,0), \\ 0 & |u^k_{i+r,j+s} - u^k_{i,j}| \geq \theta \text{ and } (r,s) \neq (0,0), \\ 1 - \sum\limits_{\substack{r,s=-1 \\ (r,s)\neq 0}}^{1} (\omega_1)^k_{i+r,j+s}\,(\omega_2)^k_{i+r,j+s} & (r,s) = (0,0), \end{cases}$$

and with periodic boundary conditions. While the adaptive smoothing procedure leads for $k \to \infty$ to a constant image with the same average grey level as the initial image, our scheme (2.2) yields an image with different constants corresponding to the obtained final partition. This property is due to the fact that in contrast to the schemes above (see [2, 11, 28, 32]) our weight measuring the radiometric distance ω_2 completely vanishes if the difference between neighboring pixel values is too big.

3.2. *Nonlinear diffusion filtering*

Consider the diffusion process

$$u_t = \operatorname{div}\left(g(|\nabla u|)\,\nabla u\right) \qquad \text{on } (0,N_1)\times(0,N_2)\times(0,\infty) \tag{3.3}$$

with the given noisy image u^0 as initial state

$$u(x,0) = u^0(x), \qquad x \in [0,N_1)\times[0,N_2),$$

and with periodic boundary conditions. Here subscripts denote partial derivatives. The time t is a scale parameter. Increasing t corresponds to stronger filtering. The divergence expression can be decomposed by means of two orthonormal basis vectors x_1, x_2,

$$\operatorname{div}\left(g(|\nabla u|)\,\nabla u\right) = \partial_{x_1}\left(g(|\nabla u|)\,\partial_{x_1} u\right) + \partial_{x_2}\left(g(|\nabla u|)\,\partial_{x_2} u\right).$$

We slightly change the diffusion equation (3.3) and consider a discretization of

$$u_t = \partial_{x_1}\left(g(|\partial_{x_1} u|)\,\partial_{x_1} u\right) + \partial_{x_2}\left(g(|\partial_{x_2} u|)\,\partial_{x_2} u\right) \quad \text{on } (0,N_1)\times(0,N_2)\times(0,\infty)$$

(see also [23]). Choosing $x_1 = (1,0)$, $x_2 = (0,1)$, and replacing the derivatives by finite differences, we obtain the discrete scheme

$$\begin{aligned}\tfrac{u_{i,j}^{k+1}-u_{i,j}^k}{\tau} &= g(|u_{i+1,j}^k - u_{i,j}^k|)\,(u_{i+1,j}^k - u_{i,j}^k) - g(|u_{i,j}^k - u_{i-1,j}^k|)\,(u_{i,j}^k - u_{i-1,j}^k)\\ &\quad + g(|u_{i,j+1}^k - u_{i,j}^k|)\,(u_{i,j+1}^k - u_{i,j}^k) - g(|u_{i,j}^k - u_{i,j-1}^k|)\,(u_{i,j}^k - u_{i,j-1}^k).\end{aligned}$$

Analogously, choosing the diagonal directions $x_1 = \frac{1}{\sqrt{2}}(1,1)$, $x_2 = \frac{1}{\sqrt{2}}(1,-1)$ we find

$$\begin{aligned}&\frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau} =\\ &= \tfrac{1}{2}\left[g\left(\tfrac{|u_{i+1,j+1}^k - u_{i,j}^k|}{\sqrt{2}}\right)(u_{i+1,j+1}^k - u_{i,j}^k) - g\left(\tfrac{|u_{i,j}^k - u_{i-1,j-1}^k|}{\sqrt{2}}\right)(u_{i,j}^k - u_{i-1,j-1}^k)\right.\\ &\quad \left. + g\left(\tfrac{|u_{i+1,j-1}^k - u_{i,j}^k|}{\sqrt{2}}\right)(u_{i+1,j-1}^k - u_{i,j}^k) - g\left(\tfrac{|u_{i-1,j+1}^k - u_{i,j}^k|}{\sqrt{2}}\right)(u_{i-1,j+1}^k - u_{i,j}^k)\right].\end{aligned}$$

Averaging the two equations leads to

$$u_{i,j}^{k+1} = u_{i,j}^k + \tau \sum_{\substack{r,s=-1\\(r,s)\neq(0,0)}}^{1} \frac{g(\sqrt{2}^{1-|r|-|s|}(|u_{i+r,j+s} - u_{i,j}|))\,(u_{i+r,j+s} - u_{i,j})}{r^2+s^2}. \quad (3.4)$$

This scheme is very similar to our scheme in (2.1) if $\alpha = \tau$ and if the diffusivity g is of the form

$$g(|x|) := \begin{cases} 1 & |x| < \theta,\\ 0 & |x| \geq \theta.\end{cases}$$

However, this diffusivity function is different from the most usual diffusivity functions used for image smoothing (see e.g. [3, 10, 23, 33]). It can be derived from the robust cup-function used e.g. in [36], see also [17, 21]. It occurs in [22], there it has been derived from hard wavelet shrinkage. Its behavior is similar to the Tukey diffusivity function used in [3] which is also derived from robust statistics. Indeed, the numerical results in [3] imply that the Tukey diffusivity preserves edges better than other diffusivities which do not completely vanish for high gradients.

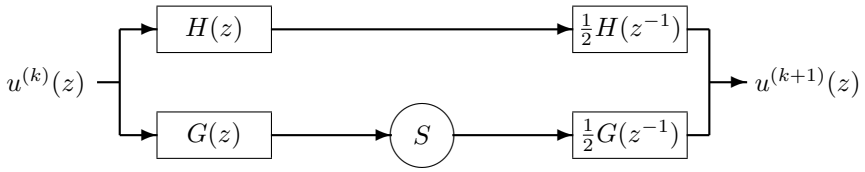


Fig. 2. Nonsubsampled two-channel filter bank with $H(z) = \frac{1+z}{\sqrt{2}}$ and $G(z) = \frac{1-z}{\sqrt{2}}$.

3.3. *Wavelet shrinkage*

We use a redundant wavelet system and apply a Haar wavelet filter bank with hard thresholding. Let us consider a translation-invariant Haar wavelet filter bank with a single scale in one dimension (see Figure 2).

Let $a^k = (a_i^k)_{i=0}^{N-1}$ be a discrete signal and $a^k(z) = \sum_{i=0}^{N-1} a_i^k z^{-i}$ its z-transform. Further, let $H(z) := \frac{1}{\sqrt{2}}(1+z)$, $G(z) = \frac{1}{\sqrt{2}}(1-z)$ the low-pass and the high-pass filter of the filter bank and S the shrinkage function. We obtain from the filter bank

$$a^{k+1}(z) = \frac{1}{2}\left[H(z^{-1})H(z)a^k(z) + G(z^{-1})S(G(z)a^k(z))\right],$$

where the shrinkage function S acts on the coefficients of $G(z)a^k(z)$, i.e.,

$$S(G(z)a^k(z)) = S\left(\sum_{i=0}^{N-1} \frac{a_i^k - a_{i+1}^k}{\sqrt{2}} z^{-i}\right) := \sum_{i=0}^{N-1} S\left(\frac{a_i^k - a_{i+1}^k}{\sqrt{2}}\right) z^{-i},$$

and where we use periodic boundary conditions. By

$$H(z^{-1})H(z) + G(z^{-1})G(z) = 2$$

we find

$$\begin{aligned} a^{k+1}(z) &= a^k(z) - \tfrac{1}{2}G(z^{-1})\left[G(z)a^k(z) - S(G(z)a^k(z))\right] \\ &= a^k(z) - \tfrac{1}{4}(1-z^{-1})(1-z)a^k(z) + \tfrac{1}{2\sqrt{2}}(1-z^{-1})S(\tfrac{(1-z)a^k(z)}{\sqrt{2}}), \end{aligned}$$

or componentwisely,

$$a_i^{k+1} = \tfrac{1}{2}a_i^k + \tfrac{1}{4}a_{i+1}^k + \tfrac{1}{4}a_{i-1}^k + \tfrac{1}{2\sqrt{2}}\left(S\left(\tfrac{a_i^k - a_{i+1}^k}{\sqrt{2}}\right) - S\left(\tfrac{a_{i-1}^k - a_i^k}{\sqrt{2}}\right)\right). \tag{3.5}$$

Now, let again $u^0 = (u_{i,j}^0)_{i=0,j=0}^{N_1-1,N_2-1}$ be the observed noisy image and apply the above wavelet shrinkage scheme (3.5) to u^k in x-direction, y-direction and the two diagonal directions. Averaging these four approximations with the weights 1/3, 1/3, 1/6, 1/6 leads to

$$u_{i,j}^{k+1} = \frac{1}{2}u_{i,j}^k + \frac{1}{12}\sum_{\substack{r,s=-1 \\ (r,s)\neq(0,0)}}^{1} \frac{u_{i+r,j+s}^k}{r^2+s^2} + \frac{1}{6\sqrt{2}}\sum_{\substack{r,s=-1 \\ (r,s)\neq(0,0)}}^{1} \frac{1}{r^2+s^2}\, S\left(\frac{u_{i,j}^k - u_{i+r,j+s}^k}{\sqrt{2}}\right),$$

where we have used that $-S(x) = S(-x)$, i.e., that S is an odd function. By $\sum^1_{\substack{r,s=-1\\(r,s)\neq(0,0)}} (r^2+s^2)^{-1} = 6$, this equation can be written as

$$u^{k+1}_{i,j} = u^k_{i,j} + \frac{1}{12} \sum^1_{\substack{r,s=-1\\(r,s)\neq(0,0)}} \frac{(u^k_{i+r,j+s} - u^k_{i,j}) - \sqrt{2}S(\sqrt{2}^{-1}(u^k_{i+r,j+s} - u^k_{i,j}))}{r^2+s^2}.$$

Comparison with (2.1) and (2.2) yields equivalence with our special scheme (2.2) for $\alpha = 1/12$ and with the hard thresholding function

$$S(x) = \begin{cases} x & |x| \geq \theta/\sqrt{2}, \\ 0 & |x| < \theta/\sqrt{2}. \end{cases}$$

4. The algorithm and numerical examples

4.1. *The algorithm*

The complete algorithm for denoising of piecewise constant images now consists of three steps. Let $u^0 = (u^0_{i,j})^{N_1-1,N_2-1}_{i=1,j=1}$ be the given noisy image

1. For a fixed thresholding parameter $\theta > 0$, a smoothing parameter $\alpha \leq \frac{1}{6}$ and a fixed number $K \in \mathbb{N}$ perform the iteration
$$u^{k+1}_{i,j} = u^k_{i,j} + \alpha \sum^1_{\substack{r,s=-1\\(r,s)\neq(0,0)}} \frac{T_\theta(u^k_{i+r,j+s} - u^k_{i,j})}{r^2+s^2}$$
for $k = 0, 1, 2, \ldots, K-1$ and using periodic boundary conditions.
2. Apply the following mean value procedure. Establish a new iteration matrix $\mathbf{A}_K$ as in (2.3) with a suitable shrinkage parameter θ_1. Take the partition $(S^K_m)^{J_K}_{m=1}$, determined by this iteration matrix $\mathbf{A}_K$, compute for $m = 1, \ldots, J_K$ the mean values
$$\mu^K_m := \frac{1}{\#S^K_m} \sum_{n \in S^K_m} u^K_n,$$
and replace each value u^K_n with $n \in S^K_m$ by μ^K_m.
3. Using the partition $(S^K_m)^{J_K}_{m=1}$ we apply the following median value procedure. All pixel values belonging to subsets of the partition with less than 6 components are replaced by the median of their neighbor pixel values.

We shortly want to discuss the choice of parameters. Let c be the contrast of the original piecewise constant image f, i.e.

$$c = \min\{|f_{i,j} - f_{i',j'}|,\ (i',j') \in N(i,j),\ f_{i,j} \neq f(i',j')\}.$$

Then a good choice for the shrinkage parameter is $\theta \approx 0.45c$. In our numerical results we have seen that this parameter should be taken independently of the signal-to-noise ratio. However, if the ratio between contrast c of the original image and the deviation σ of the added noise is small, i.e. $c/\sigma \leq 1.5$, then we suggest to

do the first iteration step in the scheme (2.2) with $\theta = \infty$. Such a "smoothing step" has been used also in the algorithm by Polzehl and Spokoiny for finding an initial estimate for image values in case of small ratios c/σ (see [27]).

As we have also seen from the examples, there is no reason to choose the smoothing parameter α very small. A smoothing parameter $\alpha \in [1/10, 1/6]$ ensures a fast smoothing process while the convergence result holds.

The optimal number K of iterations depends on the considered image and on the signal-to-noise ratio. In practice, only a few iterations of the iteration process are necessary until a partition of pixels is obtained, which does not change or only slowly changes after further iterations. One can also take an adaptive choice of K by comparing two successive iteration matrices or their corresponding partitions.

If we have arrived at the final partition already after finishing the iteration, we can choose $\theta_1 = \theta$ in order to perform the piecewise constant image which also would have been obtained, if we had performed infinitely many iterations in step 1. If the separation process has not been finished completely after the first step, our practical observations suggest to apply the shrinkage parameter $\theta_1 = \frac{\theta c}{10\sigma}$.

The third step of the algorithm improves the result especially for smaller signal-to-noise ratios.

4.2. *Numerical examples*

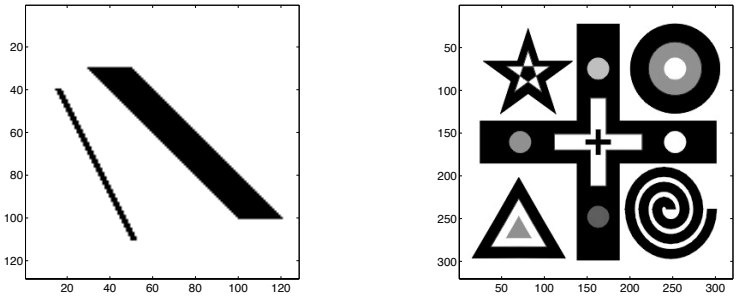
In this section, we want to show the performance of our method for denoising piecewise constant images and compare it with other methods. In particular, we consider the digital total variation filter by Chan, Osher and Shen [9], the statistical adaptive weights smoothing of Polzehl and Spokoiny [27] for piecewise constant images, the four-pixel scheme of Welk, Steidl and Weickert [35], which is a stable iterative nonlinear diffusion scheme with regularized TV diffusivity, and the noniterative method of non-local means by Buades, Coll and Morel [4, 5].

In the first test, we consider the performance of the proposed methods for denoising an artificial piecewise constant (128×128) image with a deep narrow scratch and a deep wide scratch (see Figure 3(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 0.

Fig. 4(a) shows the image contaminated with heavy noise and below a cut through the image, comparing the noisy image with the original in Fig. 3(a). In this example the ratio between contrast c and deviation σ is $c/\sigma = 3.3$, i.e., it is rather big. We applied the above mentioned denoising methods, where everytimes we tried to optimize the parameters in order to get an optimal result for each



(a) (b)

Fig. 3. Piecewise constant images with two gray levels (a) and with five gray levels (b)

method. We present the denoised image and a cut through the image (below) for all five methods to be compared.

Fig. 4(b) is obtained using 50 iterations of digital TV-filter with optimized fitting parameter $\lambda = 0.9$ (see [9]). Fig. 4(c) shows the result of the adaptive weights smoothing where exact knowledge of the contrast c and the variance of the noise σ^2 has been used. The further parameters $N_0 = 1$, $\lambda = 1.0$, $\eta = 3.0$ and 3 iterations have been taken for optimizing the result of the procedure. See [27] for an extensive description of these parameters.

Fig. 4(d) is obtained using the four-pixel scheme with a time step $\tau = 0.1$ and 30 iterations (see [35]). Fig. 4(e) shows the result using the non-local means method with $h = 12\sigma$ as proposed in [5]. Because, this last method is much more time-consuming than the others, we have not been able to optimize the parameter for the special image. Finally, Fig. 4(f) shows the result of our algorithm using 10 iterations of the iteration scheme with $\theta = 4.4$, $\alpha = 0.1$ and $\theta_1 = \theta/5$.

We observe that the methods considered in 4(b), 4(c), 4(d) and 4(f) well detect the occurrence of the thin scratch. However, the hight of the scratch tends to be smoothed out as one can see in the corresponding cuts through the images in 4(b), 4(c), 4(d) and 4(e). The scratch will be completely smoothed away taking more iterations in 4(c) end 4(d). In particular, our simple method 4(f) works very well for edge-preserving denoising. The result of the non-local means method 4(e) is less convincing than the others, such that one may say that this procedure is not really suitable for piecewise constant images.

In the second test, we show the performance for a more sophisticated (almost) piecewise constant 320×320 image (see Figure 3(b)). This image has five different gray levels. We added Gaussian noise such that we have the ratio $\rho = 1.5$, this implies here an SNR of 3.20 for the noisy image. We proceed as before using (almost) optimized parameters in each method (up to that of non-local means). Again Fig. 5(a) shows the image contaminated with heavy noise and a cut through the noisy image. In this example the ratio between contrast c and deviation σ is $c/\sigma = 0.8$.

Fig. 5(b) is obtained using 10 iterations of digital TV-filter with optimized fitting

parameter $\lambda = 7.0$ (see [9]). Fig. 5(c) shows the result of the adaptive weights smoothing where exact knowledge of the contrast c, the variance of the noise σ^2 as well as $N_0 = 9$, $\lambda = 2.5$, $\eta = 3.0$ and 4 iterations have been used (see [27]). According to the low ratio between contrast and deviation of the noise, the initial estimate of the image values has been obtained by local mean of 9 intensities as proposed in [27]. Fig. 5(d) is obtained using the four-pixel scheme [35] with a time step $\tau = 0.01$ and 35 iterations. Fig. 5(e) shows the result using the non-local means method with $h = 12\sigma$ as proposed in [5]. Finally, Fig. 5(f) shows the result of our algorithm using 50 iterations of the iteration scheme with $\theta = 0.13$, $\alpha = 0.1$ and $\theta_1 = \theta/10$. Because of the low ratio c/σ, one iteration step with $\theta = \infty$ has been used first.

We obtain satisfactory results for the methods in 5(b), 5(c), 5(d) and 5(f). It should be noted that a good result is already obtained in this experiment using only the first step of our algorithm.

5. Conclusion

A fast and efficient nonlinear algorithm has been proposed for edge-preserving denoising of piecewise constant images. We have provided a detailed mathematical analysis of the behavior of the scheme using the representation of one iteration step as a matrix vector multiplication. The iteration matrices determine a partition of the image into different regions in which it is smoothed by averaging the grey values of the region. Furthermore, the relations of the iteration scheme to average and bilateral filtering, nonlinear diffusion and wavelet shrinkage have been explored.

Moreover, convergence of the discrete diffusion schemes with popular diffusivities has been indirectly demonstrated. Good performance has been shown in numerical experiments in comparison to some existing methods. This will be significant for ongoing applications in metrology industry where we need to detect the features with sharp edges from measuring engineering surfaces (see e.g. [19]). In order to prove the convergence, we have just considered T_θ as a hard thresholding function here. More complex formulations will be designed for general images in future work. Another important extension is to apply the nonlinear scheme for image segmentation, because the iteration process of this algorithm corresponds to a partition of the image into subdomains.

References

1. G. Aubert and L. Vese, A variational method for image recovery, *SIAM J. Numer. Anal.* **34**(5) (1997) 1948–1979.
2. D. Barash, A fundamental relationship between bilateral filtering, adaptive smoothing and nonlinear diffusion equation, *IEEE Trans. on Pattern Analysis and Machine intelligence* **24**(6) (2002) 844–847.
3. M.J. Black, G. Sapiro, D. Marimont, and D. Heeger, Robust anisotropic diffusion, *IEEE Trans. on Image Processing* **7**(3) (1998) 421–432.

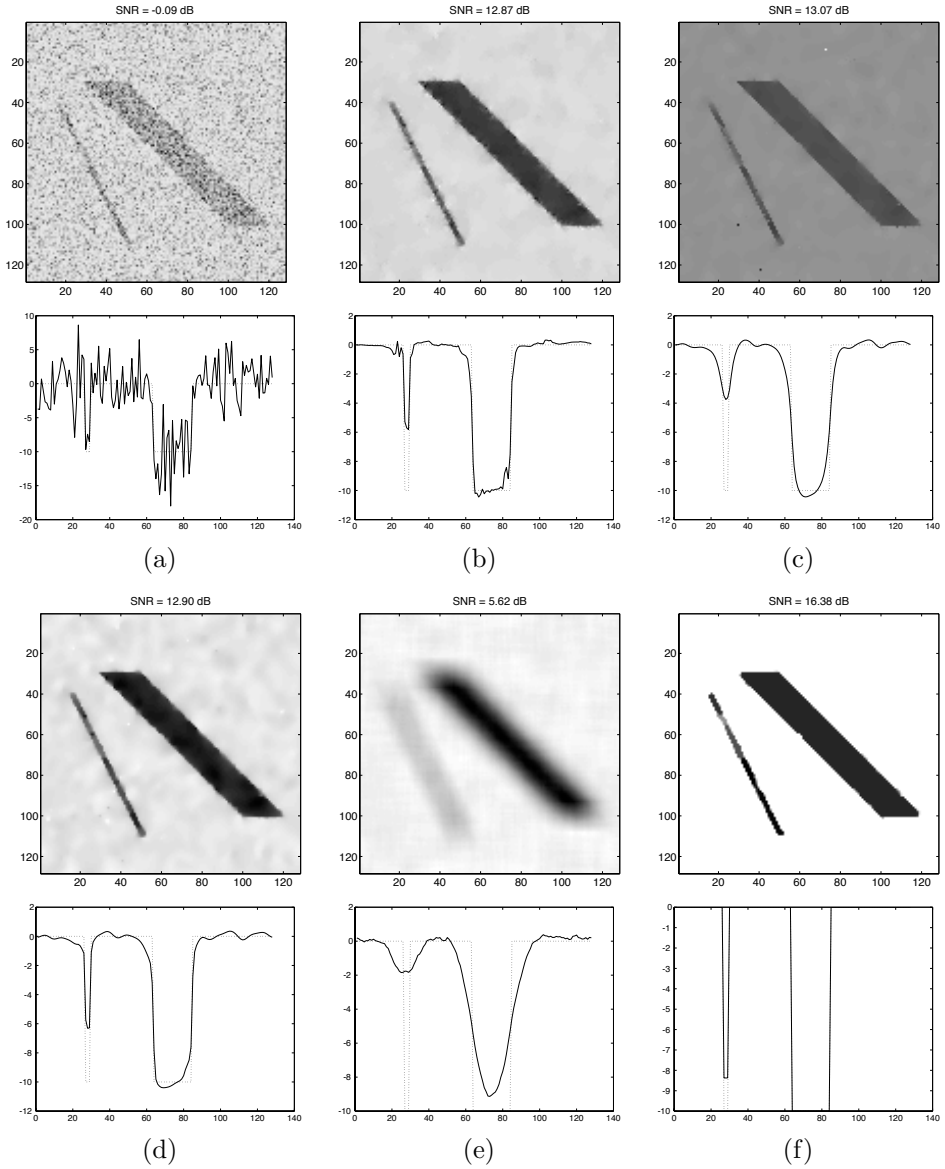


Fig. 4. Denoising of a piecewise constant image with a deep narrow scratch and a deep wide scratch.

4. A. Buades, B. Coll, and J.-M. Morel, A non-local algorithm for image denoising, *Computer Vision and Pattern Recognition* **2** (2005) 60–65.
5. A. Buades, B. Coll, and J.-M. Morel, A review of image denoising algorithms, with a new one, *Multiscale Model. Simul.* **4** (2) (2005) 490–530.
6. F. Catté, P. L. Lions, J. M. Morel and T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* **29**(1) (1992) 182–193.

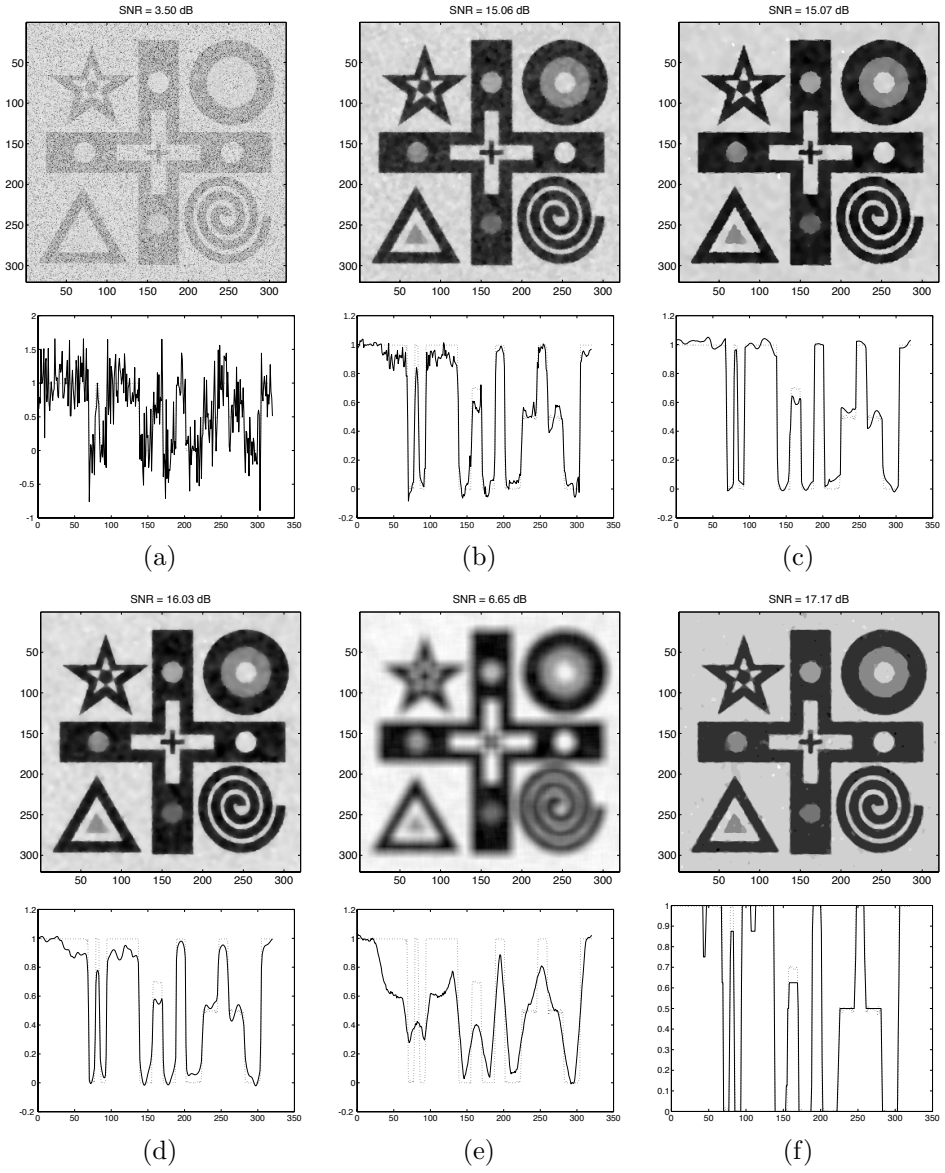


Fig. 5. Denoising of an (almost) piecewise constant image with 5 different gray levels.

7. A. Chambolle, R.A. DeVore, N. Lee, and B.L. Lucier, Nonlinear wavelet image processing: variational problems compression and noise removal through wavelet shrinkage, *IEEE Trans. on Image Processing* **7**(3) (1998) 319–335.
8. A. Chambolle and B.L. Lucier, Interpreting translationally-invariant wavelet shrinkage as a new image smoothing scale space, *IEEE Trans. on Image Processing* **10**(7) (2001) 993–1000.
9. T.F. Chan, S. Osher, and J. Shen, The digital TV filter and nonlinear denoising, *IEEE*

Trans. on Image Processing **10**(2) (2001) 231–241.
10. P. Charbonnier, 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, Austin, IEEE Computer Society Press, 1994, 168–172.
11. K. Chen, Adaptive smoothing via contextual and local discontinuities, *IEEE Trans. on Pattern Analysis and Machine intelligence* **27**(10) (2005) 1552–1567.
12. R.R. Coifman and D. Donoho, Translation invariant denoising, In: A. Antoniadis and G. Oppenheim, editors, *Wavelets in Statistics*, Springer, New York, 1995, 125–150.
13. R.R. Coifman, A. Sowa, Combining the calculus of variations and wavelets for image enhancement, *Appl. Comput. Harmon. Anal.* **9** (2000) 1–18.
14. S. Durand, J. Froment, Reconstruction of wavelet coefficients using total variation minimization, *SIAM J. Sci. Comput.* **24** (2003), 1754–1767.
15. M. Elad, On the origin of the bilateral filter and ways to improve it, *IEEE Trans. on Image Processing* **11(10)** (2002) 1141–1151.
16. L.D. Griffin, Mean, median and mode filtering of images, *Proceedings of the Royal Society of London, Series A*, **456** (2004) 2995–3004.
17. F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel, *Robust Statistics: An Approach Based on Influence Functions* (Wiley, New York, 1986).
18. R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985).
19. J. Ma, M. Fenn, Combined complex ridgelet shrinkage and total variation minimization, *SIAM J. Sci. Comput.* **28**(3) (2006), 984–1000.
20. P. Mrázek and J. Weickert, Rotationally invariant wavelet shrinkage. In *Pattern Recognition*, B. Michaelis and G. Krell (Eds.), LNCS, Springer, Berlin, 2005, vol. 2781, pp. 156–163.
21. P. Mrázek, J. Weickert, and A. Bruhn, On robust estimation and smoothing with spatial kernels, *Geometric Properties from Incomplete Data*, R. Klette, R. Kozera, L. Noakes, J. Weickert (Eds.), Springer, Dordrecht, 2005, pp. 335–352.
22. P. Mrázek, J. Weickert, and G. Steidl, Diffusion inspired shrinkage functions and stability results for wavelet denoising, *International Journal of Computer Vision* **64** (2005) 171–186.
23. P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **12** (1990) 629–639.
24. M. Nikolova, Local strong homogeneity of a regularized estimator, *SIAM J. Appl. Math.* **61**(2) (2000) 633–658.
25. G. Plonka and G. Steidl, A multiscale wavelet-inspired scheme for nonlinear diffusion, *International Journal of Wavelets, Multiresolution and Information Processing* **4** (2006) 1–22.
26. I. Pollak, A.S. Willsky, and H. Krim, Image segmentation and edge enhancement with stabilized inverse diffusion equations, *IEEE Trans. on Image Processing* **9**(2) (2000) 256–266.
27. J. Polzehl and V.G. Spokoiny, Adaptive weights smoothing with application to image segmentation, *J. R. Statist. Soc. B* **62**, Part 2, (2000) 335–354.
28. P. Saint-Marc, J.S. Chen, and G. Medioni, Adaptive smoothing: A general tool for early vision, *IEEE Trans. on Pattern Analysis and Machine intelligence* **13(6)** (1991) 514–529.
29. O. Scherzer and J. Weickert, Relations between regularization and diffusion filtering, *Journal of Mathematical Imaging and Vision* **12**(1) (2000), 43–63.
30. J. L. Starck, E. J. Candès, D. L. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11** (2002) 670–684.

31. G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk, On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides, *SIAM J. Numer. Anal.*, **42**(2) (2004), 686–713.
32. C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, *Proceedings of the 6th International Conference on Computer Vision*, Bombay, India, 1998, pp. 839–846.
33. J. Weickert, *Anisotropic Diffusion in Image Processing* (Teubner, Stuttgart, 1998).
34. J. Weickert, B.M. ter Haar Romeny, and M. A. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Trans. Image Process.* **7**(3) (1998) 398–410.
35. M. Welk, G. Steidl, and 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.* Lecture Notes in Computer Science, Springer, Berlin, 610–621 (2005).
36. G. Winkler, V. Aurich, K. Hahn and A. Martin, Noise reduction in images: Some recent edge-preserving methods, *Pattern Recognition and Image Analysis* **9**(4) (1999) 749–766.
37. Y. You, W. Xu, A. Tannenbaum, Behavioral analysis of anisotropic diffusion in image processing, *IEEE Trans. Image Process.* **5** (1996) 1539–1553.