

Combined Curvelet Shrinkage and Nonlinear Anisotropic Diffusion

JIANWEI MA¹ AND GERLIND PLONKA²

¹ Laboratoire LMC-IMAG, University Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France

jianwei.ma@imag.fr

² Department of Mathematics, University of Duisburg-Essen, Campus Duisburg, 47048

Duisburg, Germany

plonka@math.uni-duisburg.de

Abstract

In this paper, a diffusion-based curvelet shrinkage is proposed for discontinuity-preserving denoising using a combination of a new tight frame of curvelets with a nonlinear diffusion scheme. In order to suppress the pseudo-Gibbs and curvelet-like artifacts, the conventional shrinkage results are further processed by a projected total variation diffusion, in which only the insignificant curvelet coefficients or high-frequency part of the signal are changed by use of a constrained projection. Numerical experiments from piecewise-smooth to textured images show good performances of the proposed method to recover the shape of edges and important detailed components, in comparison to some existing methods.

Mathematics Subject Classification 2000. 65T60, 65M06, 65M12, 94A12.

Key words. Curvelets, nonlinear diffusion, regularization, discontinuity-preserving, denoising.

1 Introduction

Image denoising is one important operation in image processing. The tools for attacking this problem come from very different fields like Computational Harmonic Analysis (CHA) and partial differential equations (PDEs). They serve the same denoising purpose; removing the noise from the observed signal without sacrificing important discontinuity structures such as edges. Generally, tools from the CHA such as wavelets have promising properties for singularity analysis and efficient computational complexity, but suffer from the pseudo-Gibbs artifacts and shift/rotation variance. New approaches like complex wavelets and curvelets attempt to overcome these problems at least partially.

Application of PDEs such as nonlinear diffusion, variational methods and level sets are almost free from the above lacks of CHA, but cost heavy computational burden that is not suitable for time-critical application. The two directions have been opposed for long time, but recent research is increasingly focusing on combinations of both. Total variation (TV) regularization has been combined with wavelets to reduce the pseudo-Gibbs artifacts resulted from wavelet shrinkage [4, 6, 9, 13]. The relations between wavelet shrinkage, nonlinear diffusion, and regularization have been explored [19, 21, 27]. This connection provides a fruitful exchange of ideas between the two directions. Recently, diffusion-inspired wavelet shrinkage with improved rotation invariance has been considered [14]. On the other hand, a new wavelet-inspired explicit scheme that permits larger time steps while preserving convergence has been also proposed for nonlinear diffusion [18].

Within the last years, Candès and Donoho [1, 2, 3] developed a new geometric multiscale transform, the so-called curvelet transform, which allows an optimal sparse representation of objects with C^2 -singularities. The needle-shape elements of this transform own very high directional sensitivity and anisotropy. The transform represents edges and singularities along curves much more efficiently than traditional wavelet transforms. For a smooth object f with discontinuities along smooth curves, the best m -term approximation \tilde{f}_m by curvelet thresholding 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 is superior to tensor-product wavelet transforms in fields of image processing [12, 20].

The idea of image smoothing by nonlinear diffusion can be explained as follows. Let u_0 be an observed noisy image which is known to be the sum of the original image f and some Gaussian noise n ,

$$u_0(x) = f(x) + n(x), \quad x = (x_1, x_2) \in \Omega.$$

Here $\Omega \subset \mathbb{R}^2$ denotes a rectangle. We consider the diffusion process (see [16])

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla u|)\nabla u) \quad (1.1)$$

with the given noisy signal u_0 as initial condition

$$u(x, 0) = u_0(x), \quad x \in \Omega \quad (1.2)$$

and periodic boundary conditions. Here the time t acts as a scale parameter for filtering. The choice of the diffusion function $g(x) = \text{const}$ corresponds to linear diffusion which is known to lead to a strong smoothing of u with increasing t . 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.

Frequently applied diffusivities in (1.1) are Perona-Malik diffusivity $g(x) := 1/(1 + x^2/\gamma^2)$, Charbonnier diffusivity $g(x) := 1/\sqrt{1 + x^2/\gamma^2}$, and TV diffusivity $g(x) := 1/x$ with suitable chosen contrast parameter γ (see [25]). In spite of having many desirable theoretical and computational properties, one serious problem of the diffusion model is that it is very sensitive to noise (see [28]). The noise often introduces very large oscillations of the gradient ∇u , therefore the gradient-based model in (1.1) possibly misconstrues the true edges and heavy noise, which leads to undesirable diffusion in regions where there is no true edge. Another problem is that staircasing effects arise around smooth edges [25]. A remedy to the deficiencies is suggested by introducing a regularization of the model (see [5, 23]). A typical improved model is based on Gaussian regularization

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla(G_\sigma \star u)|)\nabla u) \quad (1.3)$$

where G_σ is a Gaussian kernel with variance σ . The Gaussian filtering acts as a pre-processing to reduce the influence of noise during the diffusion process. Other possibilities, e.g., wavelet regularization [24], time-delay regularization [7], nonlinear perception regularization [11], and regularization using special numerical schemes [10, 22, 27], will not be considered here.

The goal of this paper is to propose a model which is a hybrid method using curvelet shrinkage and nonlinear diffusion for discontinuity-preserving denoising. In section 2, we introduce the curvelet transform and derive an algorithm for the efficient computation of the discrete curvelet coefficients in the periodic case. In section 3, we propose the new hybrid model for image smoothing. In this model we first apply a curvelet shrinkage method and afterwards a diffusion process in order to reduce the pseudo-Gibbs phenomenon. This idea is closely related to approaches in [9, 13]. We also compare our approach with the model obtained by curvelet shrinkage followed by a projected regularization method. Section 4 is devoted to numerical results for image denoising obtained with the new model.

2 Curvelet transform

The first generation of curvelets is based on block ridgelets [1]. Apart from the blocking effects, however, the application is limited, because the geometry of ridgelets is itself unclear as they are not true ridge functions in digital images. In this paper, we apply the second-generation Discrete Curvelet Transform DCuT [2, 3], which is considerably simpler to use.

Let $V(t)$ and $W(r)$ be a pair of smooth, non-negative real-valued window functions, such that V is supported on $[-1, 1]$ and W on $[\frac{1}{2}, 2]$. The windows need to satisfy the admissibility conditions

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, \quad t \in \mathbb{R}, \quad \sum_{j=-\infty}^{\infty} W^2(2^{-j}r) = 1, \quad r > 0.$$

These conditions are satisfied taking e.g. the scaled Meyer windows (see [8], p. 137)

$$V(t) = \begin{cases} 1 & |t| \leq 1/3 \\ \cos[\frac{\pi}{2}\nu(3|t|-1)] & 1/3 \leq |t| \leq 2/3 \\ 0 & \text{else} \end{cases}, \quad W(r) = \begin{cases} 1 & 5/6 \leq r \leq 4/3 \\ \cos[\frac{\pi}{2}\nu(5-6r)] & 2/3 \leq r \leq 5/6 \\ \cos[\frac{\pi}{2}\nu(3r-4)] & 4/3 \leq r \leq 5/3 \\ 0 & \text{else} \end{cases},$$

where ν is a smooth function satisfying

$$\nu(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x \geq 1 \end{cases}, \quad \nu(x) + \nu(1-x) = 1, \quad x \in \mathbb{R}.$$

For the simple case $\nu(t) = t$ the window functions $V(t)$ and $W(r)$ are plotted in Figure 1.

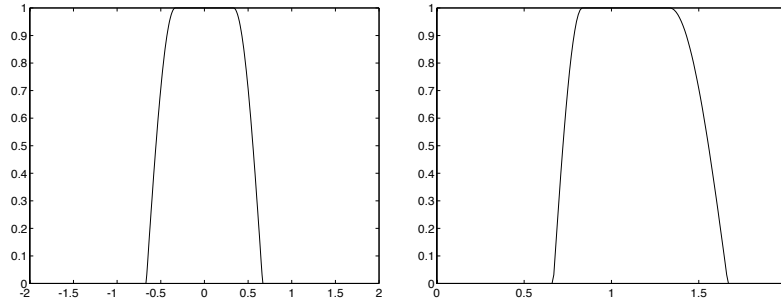


Figure 1: Plot of the windows $V(t)$ (left) and $W(r)$ (right).

Let the Fourier transform of $f \in L^2(\mathbb{R}^2)$ be defined by $\widehat{f}(\xi) := \frac{1}{2\pi} \int_{\mathbb{R}^2} f(x) e^{-i\langle x, \xi \rangle} dx$. Now, for $j \geq 0$ let the window $U_j(\xi)$, $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ in frequency domain be given by

$$U_j(\xi) = 2^{-3j/4} W(2^{-j}|\xi|) V(2^{\lfloor j/2 \rfloor} \theta), \quad \xi \in \mathbb{R}^2,$$

where $(|\xi|, \theta)$ denotes the polar coordinates corresponding to ξ . The support of U_j is a polar 'wedge' determined by $\text{supp } W(2^{-j}\cdot) = [2^{j-1}, 2^{j+1}]$ and $\text{supp } V(2^{\lfloor j/2 \rfloor}\cdot) = [-2^{-\lfloor j/2 \rfloor}, 2^{-\lfloor j/2 \rfloor}]$ (see Figure 2 for an example of the window U_0 and its support).

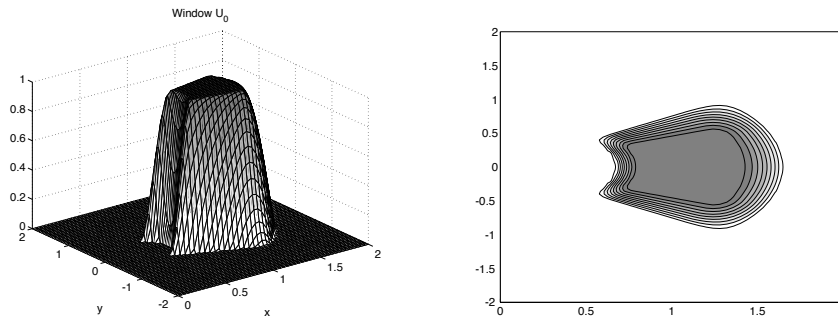


Figure 2: Window $U_0(\xi)$ (left) and its support (right).

The system of curvelets is now indexed by three parameters; a scale 2^{-j} , $j \in \mathbb{N}_0$; an equispaced sequence of rotation angles $\theta_{j,l} = 2\pi l \cdot 2^{-\lfloor j/2 \rfloor}$, $0 \leq l \leq 2^{\lfloor j/2 \rfloor} - 1$, and a position $x_k^{(j,l)} = R_{\theta_{j,l}}^{-1}(k_1 2^{-j}, k_2 2^{-\lfloor j/2 \rfloor})^T$, $(k_1, k_2) \in \mathbb{Z}^2$, where $R_{\theta_{j,l}}$ denotes the rotation matrix with angle $\theta_{j,l}$. The curvelets are defined by

$$\varphi_{j,l,k}(x) := \varphi_j(R_{\theta_{j,l}}(x - x_k^{(j,l)})), \quad x = (x_1, x_2) \in \mathbb{R}^2,$$

where $\widehat{\varphi}_j(\xi) := U_j(\xi)$, i.e., U_j is the Fourier transform of φ_j . Observe that in spatial domain, $\varphi_{j,l,k}$ is of rapid decay away from a 2^{-j} by $2^{-j/2}$ rectangle with center $x_k^{(j,l)}$ and orientation

$\theta_{j,l}$ with respect to the vertical axis in x . Further, introducing the real-valued, non-negative low-pass window W_0 by

$$W_0(r)^2 + \sum_{j \geq 0} W(2^{-j}r)^2 = 1,$$

let the coarse scale curvelet be given by

$$\varphi_{-1,0,k}(x) := \varphi_{-1}(x - k), \quad \widehat{\varphi}_{-1}(\xi) := W_0(|\xi|),$$

being nondirectional. For simplification, let $\mu = (j, l, k)$ be the collection of the triple index. The system of curvelets (φ_μ) forms a tight frame in $L^2(\mathbb{R}^2)$, i.e., each function $f \in L^2(\mathbb{R}^2)$ can be represented by

$$f = \sum_{\mu} c_{\mu}(f) \varphi_{\mu}.$$

Using Parseval's identity, the curvelet coefficients are given by

$$c_{\mu}(f) := \langle f, \varphi_{\mu} \rangle = \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{\widehat{\varphi}_{\mu}(\xi)} d\xi = \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{U_j(R_{\theta_{j,l}} \xi)} e^{i(x_k^{(j,l)}, \xi)} d\xi. \quad (2.1)$$

In practical implementations one would like to have Cartesian arrays instead of the polar tiling of frequency plane. Cartesian coroneae are based on concentric squares (instead of circles) and shears. Candès et al. [3] applied a pseudo-polar grid by replacing the window $W_j(\xi) := W(2^{-j}\xi)$ by a window of the form

$$\widetilde{W}_j(\xi) = \chi_{[0,\infty)}(\xi_1) \sqrt{\phi^2(2^{-j-1}\xi_1) - \phi^2(2^{-j}\xi_1)}, \quad j \geq 0,$$

where the one-dimensional window ϕ satisfies $0 \leq \phi \leq 1$, $\text{supp } \phi \subset [-2, 2]$ and $\phi(r) = 1$ for $r \in [-1/2, 1/2]$. (Here, $\chi_{[0,\infty)}(\xi_1)$ denotes the characteristic function of $[0, \infty)$.) As before, ϕ can be taken to be a scaled Meyer window.

With $V_j(\xi) := V(2^{\lfloor j/2 \rfloor} \xi_2/\xi_1)$ the Cartesian window,

$$\widetilde{U}_j(\xi) := 2^{-3j/4} \widetilde{W}_j(\xi) V_j(\xi)$$

can be determined, being analogous to U_j and determining the frequencies near the wedge

$$\{(\xi_1, \xi_2) : 2^j \leq \xi_1 \leq 2^{j+1}, -2^{\lfloor j/2 \rfloor} \leq \xi_2/\xi_1 \leq 2^{\lfloor j/2 \rfloor}\}.$$

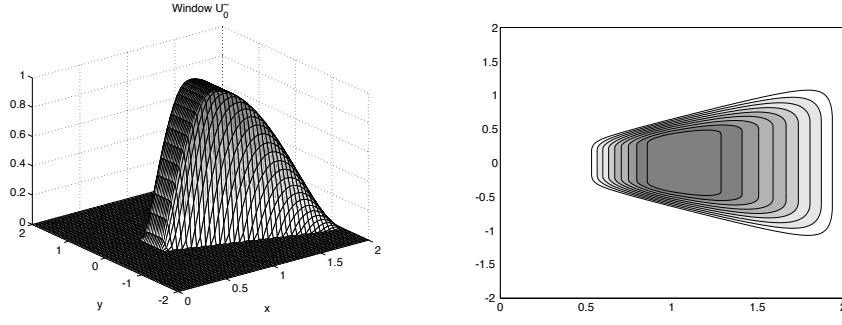


Figure 3: Window $\tilde{U}_0(\xi)$ (left) and its support (right).

See Figure 3 for an example of \tilde{U}_0 . Let $\tan \theta_{j,l} := l 2^{-\lfloor j/2 \rfloor}$, $l = -2^{\lfloor j/2 \rfloor}, \dots, 2^{\lfloor j/2 \rfloor} - 1$ be the set of equispaced slopes and put

$$\tilde{\varphi}_\mu(x) = \tilde{\varphi}_{j,l,k}(x) := \tilde{\varphi}_j(S_{\theta_{j,l}}^T(x - \tilde{x}_k^{(j,l)})), \quad x = (x_1, x_2) \in \mathbb{R}^2, \quad \widehat{\tilde{\varphi}}_j(\xi) := \tilde{U}_j(\xi)$$

as the cartesian counterpart of $\varphi_{j,l,k}$, where $\tilde{x}_k^{(j,l)} := S_{\theta_{j,l}}^{-T}(k_1 2^{-j}, k_2 2^{-\lfloor j/2 \rfloor}) =: S_{\theta_{j,l}}^{-T} k_j$, and with the shear matrix

$$S_\theta = \begin{pmatrix} 1 & 0 \\ -\tan \theta & 1 \end{pmatrix}.$$

Observe that the angles $\theta_{j,l}$ range between $-\pi/4$ and $\pi/4$ are not equispaced here, but the slopes. The set of curvelets $\tilde{\varphi}_\mu$ needs to be completed by symmetry and by rotation by $\pm\pi/2$ radians in order to obtain the whole family. We find the cartesian counterpart of the coefficients in (2.1) by

$$\tilde{c}_\mu(f) = \langle f, \tilde{\varphi}_\mu \rangle = \int_{\mathbb{R}^2} \hat{f}(\xi) \tilde{U}_j(S_{\theta_{j,l}}^{-1} \xi) e^{i\langle \tilde{x}_k^{(j,l)}, \xi \rangle} d\xi = \int_{\mathbb{R}^2} \hat{f}(S_{\theta_{j,l}} \xi) \tilde{U}_j(\xi) e^{i\langle k_j, \xi \rangle} d\xi.$$

For our application, we need the curvelet transform for functions (images) on the square. For that purpose we consider the N -periodization of $\tilde{\varphi}_\mu$,

$$\tilde{\varphi}_\mu^p(x) := \sum_{n \in \mathbb{Z}^2} \tilde{\varphi}_\mu(x - Nn), \quad x \in \mathbb{R}^2, n = (n_1, n_2),$$

where $N \in \mathbb{N}$ is fixed. The N -periodic function $f \in L^2(\Omega)$, $\Omega = [0, N]^2$, can now be written in the form

$$f = \sum_{\mu \in M} \tilde{c}_\mu^D(f) \tilde{\varphi}_\mu^p,$$

with a certain index set M which we will define below, and with

$$\tilde{c}_\mu^D(f) := \int_{\Omega} f(x) \overline{\tilde{\varphi}_\mu^p(x)} dx = \int_{\mathbb{R}^2} f(x) \overline{\tilde{\varphi}_\mu(x)} dx. \quad (2.2)$$

Observe that in the periodic case we have

$$\tilde{\varphi}_{j,l,(k_1,k_2)}^p(x) = \tilde{\varphi}_{j,l,(k_1+2^j N,k_2)}^p(x).$$

Further, assuming that l is of the form $l = \pm 2^t(2r + 1)$ with $r, t \in \mathbb{N}_0$, it follows

$$\tilde{\varphi}_{j,l,(k_1,k_2)}^p(x) = \tilde{\varphi}_{j,l,(k_1,k_2+2^{2\lfloor j/2\rfloor-t}N)}^p(x).$$

Hence, in the N -periodic case the index set M is of the form

$$\begin{aligned} M &= \{(-1, 0, (k_1, k_2)) : k_1, k_2 = 0, \dots, N - 1\} \cup \{(j, l, (k_1, k_2)) : j \in \mathbb{N}_0, \\ & \quad l = -2^{\lfloor j/2\rfloor}, \dots, 2^{\lfloor j/2\rfloor} - 1, k_1 = 0, \dots, 2^j N - 1, k_2 = 0, \dots, 2^{2\lfloor j/2\rfloor-t}N - 1\}. \end{aligned} \quad (2.3)$$

Using the Fourier series of f ,

$$f(x) = \sum_{m \in \mathbb{Z}^2} d_m(f) e^{2\pi i \langle m, x \rangle / N}, \quad d_m(f) := \frac{1}{N^2} \int_{\Omega} f(x) e^{-2\pi i \langle m, x \rangle / N} dx,$$

it follows with $k_j := (k_1 2^{-j}, k_2 2^{-j/2})$

$$\begin{aligned} \tilde{c}_\mu^D(f) &= \sum_{m \in \mathbb{Z}^2} d_m(f) \int_{\mathbb{R}^2} e^{2\pi i \langle m, x \rangle / N} \overline{\tilde{\varphi}_\mu(x)} dx \\ &= N \sum_{m \in \mathbb{Z}^2} d_m(f) \overline{\tilde{\varphi}_\mu\left(\frac{2\pi m}{N}\right)} \\ &= N \sum_{m \in \mathbb{Z}^2} d_m(f) \tilde{U}_j\left(\frac{2\pi}{N} S_{\theta_j, l}^T m\right) e^{-\frac{2\pi i}{N} \langle S_{\theta_j, l}^T m, k_j \rangle}. \end{aligned}$$

The numerical computation has been extensively described in [3]. However, the periodization of the curvelet system explained above has not been considered there. One can apply the following procedure for numerical evaluation of $\tilde{c}_\mu^D(f)$. Fix a $j_0 \in \mathbb{N}$ and compute $\tilde{c}_\mu^D(f)$ for $\mu \in M_{j_0}$, where $M_{j_0} = \{\mu \in M : j \leq j_0\}$. Analogously, compute $\tilde{c}_\mu^D(f)$ in the other three quadrants, if $\tilde{\varphi}_\mu(x_1, x_2)$ is replaced by $\tilde{\varphi}_\mu(-x_1, -x_2)$, $\tilde{\varphi}_\mu(x_2, x_1)$, and $\tilde{\varphi}_\mu(-x_2, -x_1)$, respectively.

Algorithm.

1. Apply 2D FFT to compute the Fourier coefficients $d_m(f)$ of f .
2. For all m with $S_{\theta_j, l}^T m \in \text{supp } \tilde{U}_j$ compute the product

$$d_m(f) \tilde{U}_j\left(\frac{2\pi}{N} S_{\theta_j, l}^T m\right).$$

3. Apply the inverse 2D FFT to obtain the discrete coefficients $\tilde{c}_\mu^D(f)$.

In the application, f is usually given as a discrete set, i.e. $f = (f_{i,j})_{i,j=0}^{N-1}$. For computing $d_m(f)$, one can interpolate the f by piecewise linear spline functions whose Fourier transform is known. In the last step, an unequispaced FFT can be used. The forward and inverse DCuT have the same computational cost of $\mathcal{O}(N^2 \log N)$ for an $(N \times N)$ image (see e.g. [3]).

3 Combination of curvelets with nonlinear diffusion

3.1 Description of the proposed model

Let $u_0 = (u_{0,i,j})_{i,j=0}^{N-1}$ be an observed discrete noisy signal on the square $\Omega = [0, N] \times [0, N]$ which is known to be the sum of the original image f and some Gaussian noise. We first apply curvelet shrinkage to the noisy signal u_0 , and then apply a projected nonlinear diffusion to reduce the pseudo-Gibbs and curvelet-like artifacts.

Let S be a hard thresholding function defined by a fixed threshold λ

$$S(x) = \begin{cases} x, & x \geq \lambda, \\ 0, & x < \lambda. \end{cases}$$

After applying the periodic curvelet transform and shrinkage to u_0 and simultaneously retaining the low-frequency components we obtain

$$u_c = \sum_{\mu \in M'} S(\tilde{c}_\mu^D(u_0)) \tilde{\varphi}_\mu^p + \sum_{\mu \in M \setminus M'} \tilde{c}_\mu^D(u_0) \tilde{\varphi}_\mu^p, \quad (3.1)$$

where $\tilde{c}_\mu^D(u_0)$ are the curvelet coefficients of u_0 as in (2.2) and with the index set

$$M' = M \setminus \{(-1, 0, (k_1, k_2) : k_1, k_2 = 0, \dots, N-1)\},$$

where M is given in (2.3). We collect all indices of the curvelet coefficients in the coarsest level and of those retained after hard thresholding in Λ , i.e.,

$$\Lambda := \{\mu \in M : |\tilde{c}_\mu^D(u_0)| \geq \lambda\} \cup M \setminus M'.$$

Now the following idea is applied. We want to keep the low frequency coefficients and the important curvelet coefficients (being greater than a threshold λ) almost untouched, but we

want to slightly change the curvelet coefficients which have been set to zero by the shrinkage procedure in such a way, that the image is smoothed. Therefore, we will not apply the nonlinear diffusion process directly to u_c but only to the difference image $u_0 - u_c$. After some iteration steps of the diffusion scheme, the smoothed difference image is then added to u_c in order to obtain the final result.

For that purpose, let the linear subspace $V \subset L^2(\Omega)$ be given by

$$V = V(u_0) := \{v : \tilde{c}_\mu^D(v) = 0, \forall \mu \in \Lambda\}. \quad (3.2)$$

Then a function of the form $u = u_c + v_c$ with u_c in (3.1) and $v_c \in V$ leaves the important coefficients $\tilde{c}_\mu^D(u_0)$ untouched, i.e.,

$$\tilde{c}_\mu^D(u) = \tilde{c}_\mu^D(u_0) = \tilde{c}_\mu^D(u_c) \quad \mu \in \Lambda.$$

Our model can be formulated as follows. We consider the diffusion process

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla P_V(u)|) \nabla P_V(u)) \quad (3.3)$$

with the noisy signal u_0 as initial condition

$$u(x, 0) = u_0(x), \quad x \in \Omega,$$

and with periodic boundary conditions. Here $P_V(u)$ denotes the projection of u onto V determined by curvelet shrinkage. More precisely, let T be the periodic curvelet transform and T^{-1} be the inverse transform, then we can write $P_V(u) = T^{-1} S^{-1} T(u)$, where the S^{-1} denotes a so-called inverse thresholding function, which only takes the coefficients outside Λ , i.e.,

$$S^{-1}(x) := \begin{cases} 0, & x \geq \lambda \\ x, & x < \lambda. \end{cases}$$

Discretization of (3.3) in time leads to

$$\frac{u^{j+1} - u^j}{\tau} = \nabla \cdot (g(|\nabla P_V(u^j)|) \nabla P_V(u^j))$$

with $u^0 = u_0$, where $u^0 = u_c + v_c$ with some $v_c \in V$. The time step τ is a scale parameter.

Observe that this iteration process leads to a sequence of images $\{u^j\}$, which can be also found by $u^j = u_c + v^j$, where $\{v^j\}$ is obtained using the iteration scheme

$$\frac{v^{j+1} - v^j}{\tau} = \nabla \cdot (g(|\nabla(P_V(v^j))|) \nabla P_V(v^j)) \quad (3.4)$$

with $v^0 := v_c = u_0 - u_c$ denoting the difference image after discrete curvelet transform and thresholding, and where we again use periodic boundary conditions.

This observation can be simply shown by induction. For $j = 0$ we have by $P_V(u^0) = P_V(u_c + v_c) = P_V(v_c) = v_c$ that

$$u^1 - u^0 = v^1 - v^0$$

and the assertion follows. Now, assuming that $u^j = v_c + v^j$ for some $j \in \mathbb{N}$, it follows by $P_V(u^j) = P_V(u_c + v^j) = P_V(v^j)$ that

$$u^{j+1} - u^j = v^{j+1} - v^j,$$

i.e., $u^{j+1} = v^{j+1} - v^j + u^j = v^{j+1} - v^j + (u_c + v^j) = v^{j+1} + u_c$.

The idea behind this approach is the following.

While u_c contains the important features of the image to be reconstructed, the difference image v_c particularly contains high-frequency components which mainly correspond to noise. The strategy to apply the diffusion process only to v_c avoids that the narrow peaks/textures are smoothed too much as in conventional diffusion. It can better remain the signal amplitude of detailed components while reducing the pseudo-Gibbs oscillations at the same time, in comparison to those methods directly using curvelet shrinkage or diffusion.

3.2 Numerical procedure

From numerical point of view, the treatment of the diffusion process can be strongly simplified by considering just

$$\frac{v^{j+1} - v^j}{\tau} = \nabla \cdot (g(\nabla(v^j)) \nabla(v^j)) \tag{3.5}$$

with initial condition $v^0 = v_c$ instead of equation (3.4). Indeed, since $v^0 = v_c \in V$, the difference $P_V v^j - v^j$ will be small.

We apply a few steps of the diffusion process to v_c using a suitable Euler-forward scheme for the discretization of (3.5) and then add the obtained smoothed difference image to u_c in order to get the desired reconstruction.

Let us describe the numerical procedure more explicitly.

In the first step we apply the periodic curvelet transform (as described in Section 2) to the noisy image u_0 , then perform the shrinkage procedure with a suitable threshold λ and apply the inverse curvelet transform to obtain u_c and the collection of indices Λ .

In the second step we apply the diffusion process (3.3) resp. (3.4) or (3.5) to suppress the pseudo-Gibbs and curvelet-like artifacts. In practical implementation, the low frequency coefficients, i.e., the smooth components in the coarsest scale are always kept untouched. In the diffusion equation the time t is a scale parameter. Increasing t corresponds to stronger filtering. The divergence expression in (3.5) can be decomposed by means of two orthonormal basis vectors x_1, x_2 ,

$$\nabla \cdot (g(|\nabla v|) \nabla v) = \partial_{x_1}(g(|\nabla v|) \partial_{x_1} v) + \partial_{x_2}(g(|\nabla v|) \partial_{x_2} v).$$

We slightly change the diffusion equation (3.5), as it is done also by Perona and Malik [16] and consider a discretization of

$$v_t = \partial_{x_1}(g(|\partial_{x_1} v|) \partial_{x_1} v) + \partial_{x_2}(g(|\partial_{x_2} v|) \partial_{x_2} v) \quad \text{on } \Omega \times (0, \infty).$$

Choosing $x_1 = (1, 0)$, $x_2 = (0, 1)$, and replacing the derivatives by finite differences, we obtain the discrete scheme

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

Here, $v_{i,j}^k$ denotes the sampled values of v^k , i.e., $v_{i,j}^k = v^k(i, j)$ for the suitable scaled image and $v^0 = v_c$. Analogously, choosing the diagonal directions $x_1 = \frac{1}{\sqrt{2}}(1, 1)$, $x_2 = \frac{1}{\sqrt{2}}(1, -1)$ we find

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

Averaging the two equations leads to

$$v_{i,j}^{k+1} = v_{i,j}^k + \tau \sum_{\substack{r,s=-1 \\ (r,s) \neq (0,0)}}^1 \frac{g(\sqrt{2}^{1-|r|-|s|} (|v_{i+r,j+s}^k - v_{i,j}^k|)) (v_{i+r,j+s}^k - v_{i,j}^k)}{r^2 + s^2}. \quad (3.6)$$

As diffusivity function g we propose to take the Perona-Malik diffusivity $g(|x|) = 1/(1 + x^2/\gamma^2)$. Other diffusivities (see e.g. [15]) can be used similarly. Concerning the convergence of the above

diffusion scheme we can apply ideas of Weickert [25, 26] and ideas similar to convergence results in [17].

Taking our original model using equation (3.3) resp. (3.4) (with a suitable spatial discretization by finite differences) or the simplified model (3.5) resp. (3.6), one only needs a few iterations of diffusion to get a satisfying result. Observe that for equation (3.5) the resulting smoothed difference image v^k is usually not longer completely in the subspace V such that in the final reconstruction result $u = u_c + v^k$ the significant curvelet coefficients (being greater than λ) can change slightly. However, since $v_c \in V$ contains no low frequency components, we have $\mathbf{v}^T \mathbf{1} = \mu = 0$ such that during the diffusion process, v^k converges to zero. In order to keep the significant coefficients of u_c resp. u_0 completely untouched, one can again apply a curvelet transform to v^k in order to compute the projection $P_V v^k$ of v^k with respect to the space V in (3.2) and take the reconstruction result $u = u_c + P_V v^k$. However, to prove the properties of prime model (3.3) resp. (3.4) is still a room for substantial progress in future research.

3.3 Convergence analysis of the numerical scheme

Let us have a closer look at the convergence of the simplified diffusion scheme (3.6). We choose one index pixel numbering of the difference images v^k . Put $N_1 = N^2$ and

$$n = i + Nj, \quad i = 0, \dots, N-1, \quad j = 0, \dots, N-1,$$

such that the pixel n ($0 \leq n \leq N_1$) corresponds to (i, j) . Further, let

$$g_{p,n}^k := g(|v_n^k - v_p^k|) = g(|v_{i,j}^k - v_{i',j'}^k|),$$

where $n = i + Nj$, $p = i' + Nj'$.

Then the iteration scheme (3.6) can be written in matrix-vector form as

$$\mathbf{v}^{k+1} = \mathbf{A}^k \mathbf{v}^k, \tag{3.7}$$

where $\mathbf{v}^k = (v_0^k, \dots, v_{N_1-1}^k)^T$ and where $\mathbf{A}^k = (A_{n,p}^k)_{n,p=0}^{N_1-1} \in \mathbb{R}^{N_1 \times N_1}$ is a sparse matrix with entries

$$A_{n,n}^k := 1 - \tau(g_{n,n-1}^k + g_{n,n+1}^k + g_{n,n-N}^k + g_{n,n+N}^k)$$

$$\begin{aligned}
& -\frac{\tau}{2}(g_{n,n-1+N}^k + g_{n,n-1-N}^k + g_{n,n+1+N}^k + g_{n,n+1-N}^k) \quad \text{for } n = 0, \dots, N_1 - 1, \\
A_{n,p}^k & := \tau g_{n,p}^k \quad \text{for } p \in \{n-1, n+1, n-N, n+N\} \pmod{N} \quad n = 0, \dots, N_1 - 1, \\
A_{n,p}^k & := \frac{\tau}{2} g_{n,p}^k \quad \text{for } p \in \{n-1-N, n-1+N, n+1-N, n+1+N\} \pmod{N}, \\
& n = 0, \dots, N_1 - 1.
\end{aligned}$$

Then we obtain

Theorem 3.1 *The iteration scheme (3.7) with a diffusivity $g(x)$ satisfying $0 < g(x) \leq 1$ for $x \in \mathbb{R}$ converges for every $\mathbf{v}^0 = (v_0^0, \dots, v_{N_1-1}^0)^T$ to the spatial average*

$$\lim_{k \rightarrow \infty} \mathbf{v}^k = \mu \mathbf{1},$$

if the time step τ satisfies $0 < \tau < 1/6$. Here $\mu := \frac{1}{N_1} \sum_{n=0}^{N_1-1} v_n^0$ is the average value of the difference image v_c .

Proof. Since $0 < \tau < 1/6$ and $0 < g(x) \leq 1$ for all $x \in \mathbb{R}$, we find that all entries of \mathbf{A}^k are non-negative, i.e. $\mathbf{A}^k \geq \mathbf{0}$, and moreover, all entries in the main diagonal are strictly positive. Further, the structure of \mathbf{A} directly implies that $\mathbf{A}^k \mathbf{1} = \mathbf{1}$ where $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^{N_1}$, i.e., the sum of elements in each row of \mathbf{A}^k is equal to 1. Hence, by symmetry of \mathbf{A}^k one simply derives that the spectral norm of \mathbf{A} is equal to 1. Consequently, following the proof of Weickert (see [25]), it follows that the eigenvalue 1 of \mathbf{A}^k is simple. Observing that

$$\sum_{n=0}^{N_1-1} v_n^{k+1} = \mathbf{1}^T \mathbf{v}^{k+1} = \mathbf{1}^T \mathbf{A}^k \mathbf{v}^k = \mathbf{1}^T \mathbf{v}^k = \sum_{n=0}^{N_1-1} v_n^k,$$

the assertions of the theorem follow.

q.e.d.

3.4 Comparison to TV-Minimization

It should be noted that a similar strategy has been used in TV-minimization based wavelet/ridgelet transform [9, 13]. In that work, the TV minimization does not set the insignificant wavelet/ridgelet coefficients to zero, but typically inputs optimal small values to cancel the pseudo-Gibbs oscillations.

Let us briefly compare TV-minimization with our model proposed above. We can formulate the TV-minimization for curvelet transform as follows. For a function u with $|\nabla u| \in L^1(\Omega)$, the total variation functional is defined by

$$TV(u) = \int_{\Omega} |\nabla u(x)| dx,$$

where Ω is a rectangle (or square) in \mathbb{R}^2 in our application. The idea of TV-minimization (see [9]) is to remove the pseudo Gibbs oscillations by minimizing the functional

$$F(u) = \int_{\Omega} |u - u_0|^2 dx + \lambda TV(u)$$

for $u \in \{u_c + v, v \in V\}$, where u_0 is the initial noisy signal and u_c is the reconstructed function after curvelet thresholding in (3.1), and where the space V is given in (3.2).

Using u_c as initial guess, i.e., $u^0 := u_c$, the constrained TV-minimization can be computed by a projected gradient descent scheme (see [9])

$$u^{l+1} = u^l - t_l P_V(g_{TV}(u^l)), \quad (3.8)$$

where $g_{TV}(u)$ denotes the subgradient of TV at u and P_V again denotes the projection of $g_{TV}(u^l)$ on V given in (3.2). The step size $t_l > 0$ can be taken in the form $t_l = t_0/\sqrt{l}$ in order to ensure convergence. It has been demonstrated that in the 1D case that space discrete TV diffusion and TV regularization/minimization are identical if we identify the diffusion time with the regularization parameter, see [21]. However, the constrained minimization scheme (3.8) is not equivalent to our model in (3.3). We can deduce the gradient of total variation functional as in [28],

$$g_{TV}(u^l) = -\nabla \cdot (g(|\nabla u^l|) \nabla u^l), \quad g(x) = \frac{1}{x}.$$

So (3.8) can be rewritten as

$$u^{l+1} = u^l + t_l P_V \nabla \cdot (g(|\nabla u^l|) \nabla u^l)$$

which can be seen as discretization of

$$\frac{\partial u}{\partial t} = P_V \nabla \cdot (g(|\nabla u|) \nabla u).$$

This equation clearly shows the difference to our model in (3.3), in which the diffusion is directly applied to the projected image $P_V u$.

4 Numerical experiments

In the first test, we show performances of the proposed method for denoising an piecewise-smooth image with line singularities, in comparison to conventional hard shrinkage. Fig. 4(a) shows the image contaminated with heavy noise. Fig. 4(b) is obtained using 'Db6' wavelet shrinkage. Nonsmoothness oscillations along the edges can be seen obviously due to the poor ability of wavelets at presenting line singularities. Fig. 4(c) is found using DCuT shrinkage. Although the DCuT is effective in recovering edges, it suffers from the pseudo-Gibbs and curvelet-like artifacts yet. Fig. 4(d) is the result obtained using the proposed curvelet-diffusion hybrid method with 15 iterations where the TV diffusivity is considered in this diffusion. The edges are preserved well while the artifacts are suppressed. The applied numerical scheme (see (3.6)) for nonlinear diffusion is similar to the four-pixel scheme of Welk, Steidl and Weickert in [27]. Repeatedly, Fig. 4(e) is obtained using DCuT but with a smaller threshold, and (f) is the improved result using our proposed method with 25 iterations. We can see that the embedded nonlinear diffusion synthesis really improved the performance of original DCuT. It is also relatively robust for different thresholds.

Our next test shows denoising results for a magnetic resonance image of a human head. Fig. 5(a) is a measured image. Fig. 5(b) represents the noisy image. Fig. 5(c) is obtained using DCuT. Fig. 5(d) is found using classical TV diffusion with 20 iterations. Fig. 5(e) is obtained using Gaussian-regularized PM diffusion with 50 iterations, Fig. 5(f) is the result using our proposed method. Fig. 6 shows the large close-up of hindbrain taken from Fig. 5 (c)-(f). It is clearly to see that the shape of edges is recovered well using our proposed method.

As one can see from the last example shown in Fig. 7, our method is especially very promising for textured images. Fig. 7(a) represents the noisy Barbara image. Fig. 7(b) is a denoised result obtained using Weickert's coherence-enhancing diffusion with coherence parameter 3, step size 0.1, iteration 20 times, and a Gaussian filtering (Sigma=0.5) is used before the gradient is calculated. Fig. 7(c) and (d) are obtained using TV diffusion with 0.0005 step size and 60 iterations, and the proposed curvelet-diffusion method with 0.0005 step size and 11 iterations, respectively. In the later two tests, the gray value of input image have been normalized. Fig.

7(e) and (f) show the removed components by (c) and (d). The parameters that we used for each method have been taken to optimize the SNR and visual quality of each method independently. The proposed method shows good performance for the texture-preserving denoising.

5 Conclusions

After considering the discrete curvelet transform by means of a special Meyer window functions, we proposed a hybrid method for image denoising by combining the curvelet shrinkage with a nonlinear diffusion, two quite different techniques at the first blush. The edges/textures can be preserved well taking advantage of the fact that the two techniques benefit from each other. This model can be interpreted as a diffusion-based curvelet synthesis. The signal is used as an initial value to complete a projected diffusion, in which only the insignificant curvelet coefficients (high-frequency part) of the signal are changed by use of a projection P_V . Diffusion here is used to suppress the pseudo-Gibbs oscillations.

As mentioned in section I, another important issue is to improve the ill-posed problem of PM diffusion. This motivates us to consider a curvelet-regularized nonlinear diffusion in future research. Essentially, we want to study

$$\frac{\partial u}{\partial t} = \nabla \cdot (g(|\nabla(P_\lambda \star u)|) \nabla u),$$

where, compared with formula (1.3), the Gaussian kernel G_σ is replaced by a curvelet shrinkage operator. Such a model will be addressed in our coming paper, as a part II of the combination of curvelets and nonlinear diffusion.

Acknowledgment

The authors like to thank the unknown referees for their valuable comments which led to a significant improvement of the paper.

The first author likes to thank for the support of the IDOPT project (CNRS-INRIA-UJF-INPG), as well as for the Guest Professorship at the Department of Mathematics of the University Duisburg-Essen.

References

- [1] 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.
- [2] 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.
- [3] E. J. Candès, L. Demanet, D. L. Donoho, L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.* **5** (2006), 861–899.
- [4] E. J. Candès, F. Guo, New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction, *Signal Process.* **82** (2002), 1519–1543.
- [5] F. Catté, P. L. Lions, J. M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* **29** (1992), 182–193.
- [6] R.R. Coifman, A. Sowa, Combining the calculus of variations and wavelets for image enhancement, *Appl. Comput. Harmon. Anal.* **9** (2000), 1–18.
- [7] G. H. Cottet, M. E. Ayyadi, A volterra type model for image processing, *IEEE Trans. Image Process.* **7** (3) (1998), 292–303.
- [8] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
- [9] S. Durand, J. Froment, Reconstruction of wavelet coefficients using total variation minimization, *SIAM J. Sci. Comput.* **24** (2003), 1754–1767.
- [10] F. Fontaine, S. Basu, Wavelet-based solution to anisotropic diffusion equation for edge detection, *Int. J. Imaging Sys. Tech.* **9** (1998), 356–368.
- [11] J. Gutiérrez, F. Ferri, J. Malo, Regularization operators for natural images based on nonlinear perception models, *IEEE Trans. Image Process.* **15** (1) (2006), 189–200.
- [12] 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.
- [13] J. Ma, M. Fenn, Combined complex ridgelet shrinkage and total variation minimization, *SIAM J. Sci. Comput.* **28** (3) (2006), 984–1000.

- [14] P. Mrázek, J. Weickert, Rotationally invariant wavelet shrinkage. In Pattern Recognition, B. Michaelis and G. Krell (Eds.), LNCS, Springer, Berlin, 2005, vol. 2781, pp. 156-163.
- [15] P. Mrázek, J. Weickert, G. Steidl, Diffusion inspired shrinkage functions and stability results for wavelet denoising, *Int. J. Computer Vision* **64** (2005), 171–186.
- [16] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.* **12** (1990), 629–639.
- [17] G. Plonka, J. Ma, Convergence of an iterative nonlinear scheme for denoising of piecewise constant images, *Int. J. of Wavelets, Multiresolution and Information Processing*, to appear.
- [18] G. Plonka, G. Steidl, A multiscale wavelet-inspired scheme for nonlinear diffusion, *Int. J. of Wavelets, Multiresolution and Information Processing* **4** (2006), 1–21.
- [19] O. Scherzer, Scale space methods and regularization for denoising and inverse problems, *Adv. Imaging and Electron Phys.* **128** (2003), 445–530.
- [20] J. L. Starck, E. J. Candès, D. L. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11** (2002), 670–684.
- [21] G. Steidl, J. Weickert, T. Brox, P. Mrázek, 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.
- [22] G. Teschke, M. Zhariy, J. Soares, A regularization of nonlinear diffusion equations in a multiresolution framework, preprint, July 2006, submitted.
- [23] F. Torkamani-Azar, K. E. Tait, Image recovery using anisotropic diffusion equation, *IEEE Trans. Image Process.* **5** (1996), 1573–1578.
- [24] J. Z Wang, Wavelet oriented anisotropic diffusion in image enhancement, preprint, 2004.
- [25] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.
- [26] 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.
- [27] 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).

- [28] Y. You, W. Xu, A. Tannenbaum, Behavioral analysis of anisotropic diffusion in image processing, *IEEE Trans. Image Process.* **5** (1996), 1539–1553.

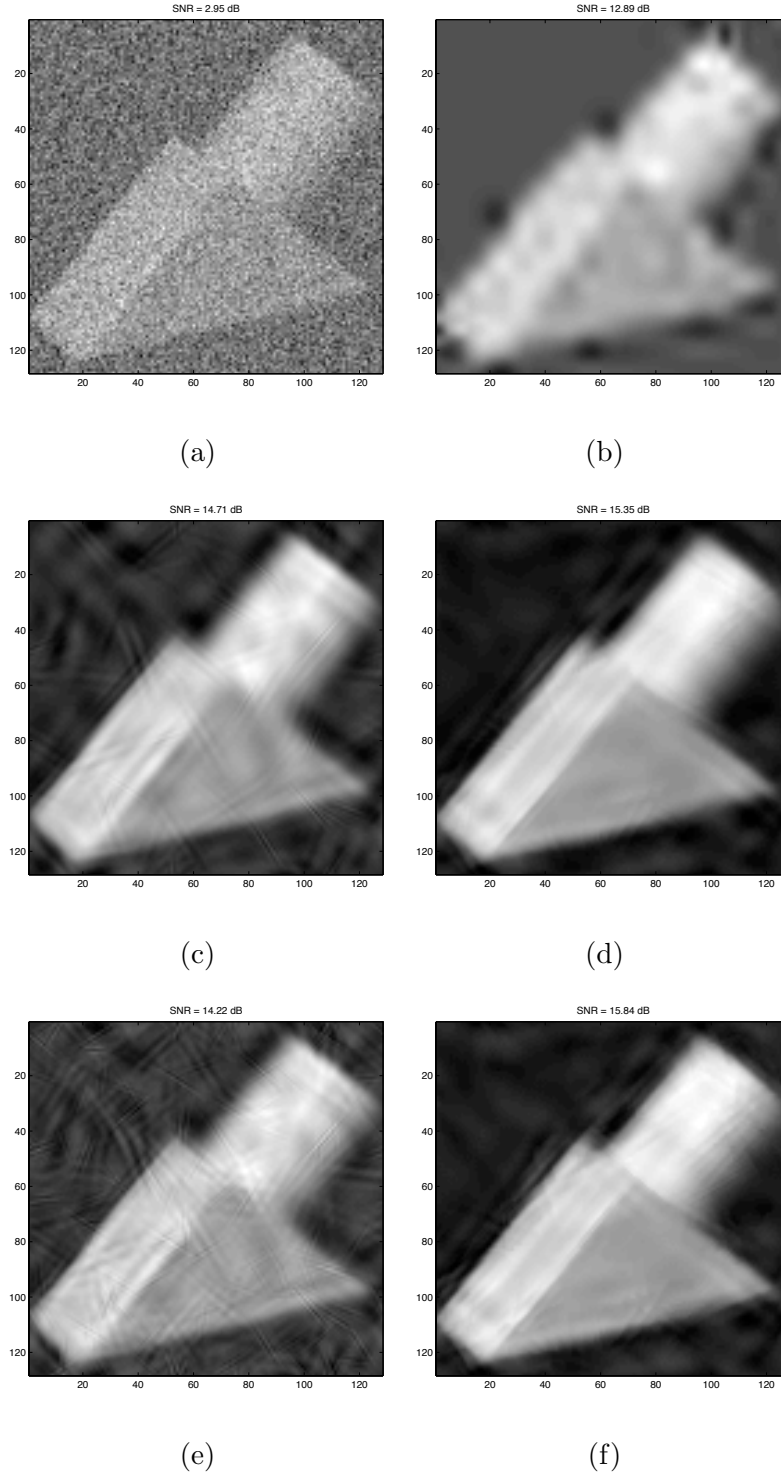


Figure 4: Denoising of a piecewise-smooth image. (a) Noisy image. (b) Denoised result by Db6 wavelet shrinkage. (c) DCuT shrinkage. (d) The proposed curvelet-diffusion method with 15 iterations. (e) DCuT with a small threshold. (f) Curvelet-diffusion method with 25 iterations.

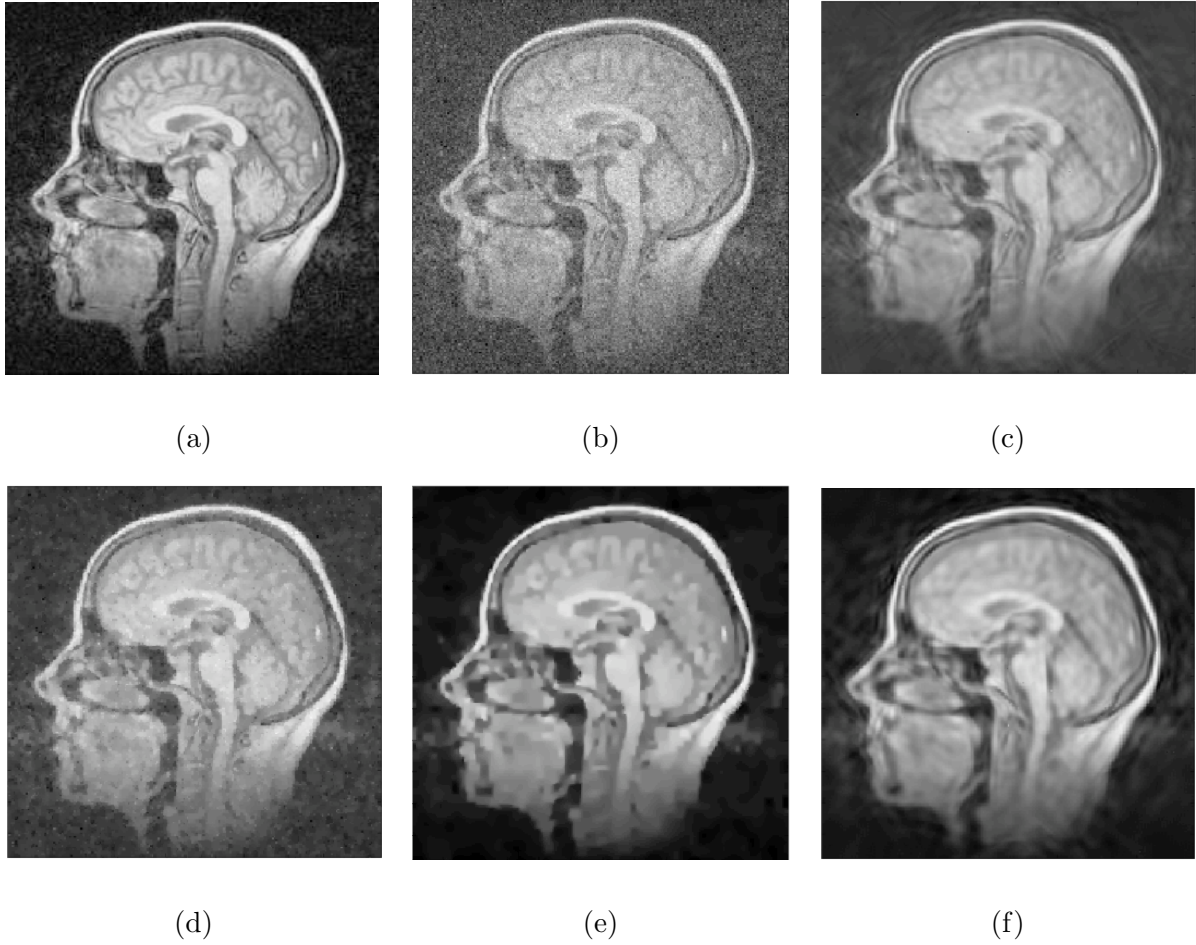


Figure 5: Denoising of a magnetic resonance image. (a) Measured image. (b) Noisy image. (c) Denoised result by DCuT shrinkage. (d) TV diffusion. (e) Gaussian-regularized PM diffusion. (f) The proposed curvelet-diffusion method.

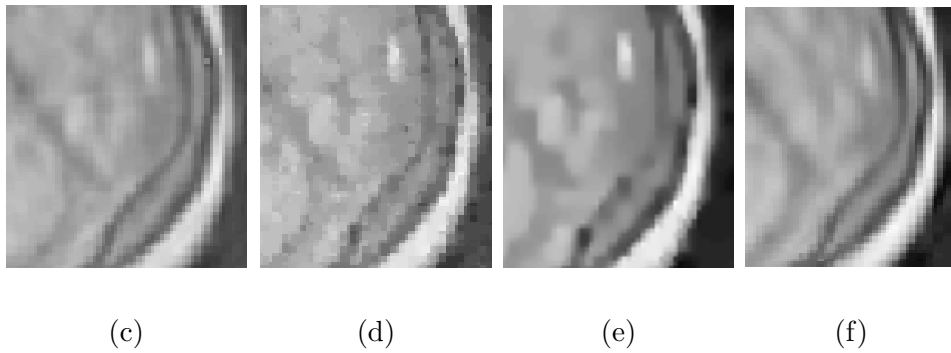


Figure 6: From left to right: Close-up of hindbrain taken from Fig. ?? (c)–(f)

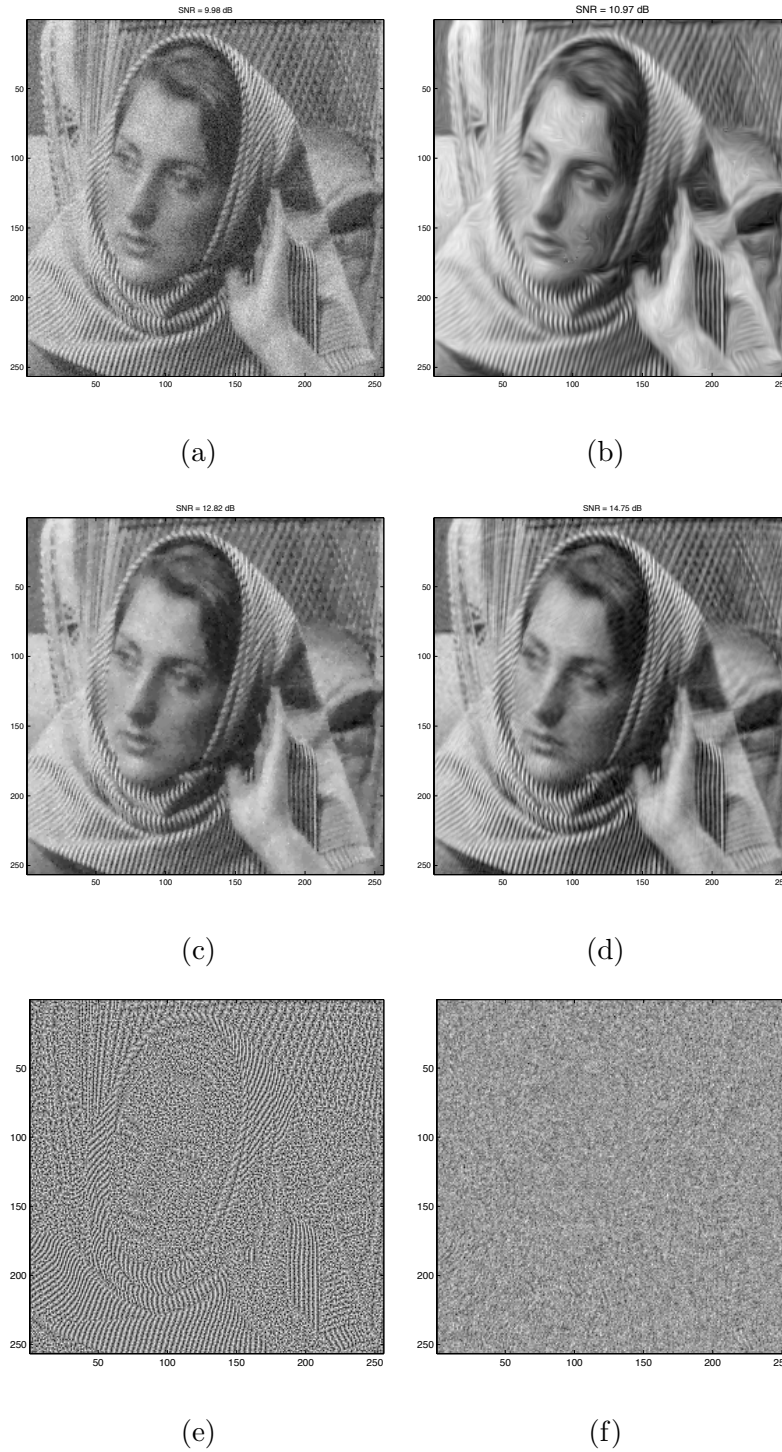


Figure 7: Denoising of an image with textures using the proposed method in comparison to nonlinear diffusion. (a) Noisy Barbara image. (b) Coherence-enhancing diffusion with 20 iterations. (c) TV diffusion with 60 iterations. (d) Our method with 11 iterations. (e) and (f) are the removed components by (c) and (d) respectively.