Weighted and well-balanced anisotropic diffusion scheme for image denoising and restoration

V. B. Surya Prasath*

Department of Computer Science, University of Missouri-Columbia, MO 65211-2060, USA

D. Vorotnikov

CMUC, Department of Mathematics, University of Coimbra, 3001–454 Coimbra, Portugal

Abstract

Anisotropic diffusion is a key concept in digital image denoising and restoration. To improve the anisotropic diffusion based schemes and to avoid the well-known drawbacks such as edge blurring and 'staircasing' artifacts, in this paper, we consider a class of weighted anisotropic diffusion partial differential equations (PDEs). By considering an adaptive parameter within the usual divergence process, we retain the powerful denoising capability of anisotropic diffusion PDE without any oscillating artifacts. Well-balanced flow version of the proposed scheme is considered which adds an adaptive fidelity term to the usual diffusion term. The scheme is general, in the sense that, different diffusion coefficient functions can be utilized according to the need and imaging modality. To illustrate the advantage of the proposed methodology, we provide some examples, which are applied in restoring noisy synthetic and real digital images. A comparison study with other anisotropic diffusion based schemes highlight the superiority of the proposed scheme.

Key words: Image restoration, Edge preserving, Nonlinear diffusion, Biased anisotropic diffusion, Well-balanced flow

---

*Corresponding author. Tel.: +1 573-882-8391 Fax: +1 573-882-8318.

Email addresses: `prasaths@missouri.edu` (V. B. Surya Prasath), `mitvorot@mat.uc.pt` (D. Vorotnikov)

Preprint submitted to Nonlinear Analysis: Real World Applications *October 15, 2013*

1. Introduction

Image denoising is one of the foremost tasks in digital image processing pipeline. There exist various methodologies for removing noise in images and the areas of image restoration and edge detection have been considered by many authors. Starting with the pioneering work of Perona and Malik [1], diffusion based partial differential equations (PDEs) are widely used in image noise removal and edge detection, see [2] for a review. Let u_0 be the noisy image which needs to be restored by removing noise without removing salient structures in it. Mathematically, $u_0 : \Omega \to \mathbb{R}$ represents a noisy version of a true image, and it is obtained by the following imaging process

$$u_0 = u + n, \tag{1}$$

here we assume that the noise process n is additive Gaussian noise with known mean and variance σ_n. The image domain $\Omega \subset \mathbb{R}^2$ is a bounded domain, usually a rectangle.

The Perona-Malik scheme (PM) can be written as a time dependent PDE, for $x \in \Omega$

$$\frac{\partial u(x,t)}{\partial t} = div\left(c\left(|\nabla u(x,t)|\right)\nabla u(x,t)\right) \tag{2}$$

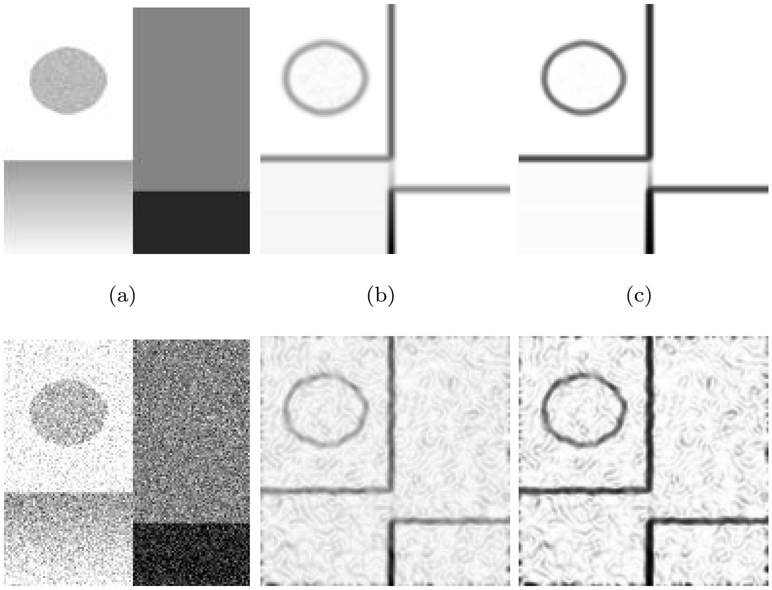
with $u(x,0) = u_0(x)$, i.e. the input noisy image is the initial datum, and the above PDE is run for a finite time $T > 0$ to obtain the denoised image $u(\cdot, T)$. The choice of the diffusion function $c : [0,\infty) \to [0,\infty)$ is important in controlling the smoothing and even enhancement of edges. In [1] the following two diffusion functions are considered

$$c_{pm1}(s) = \frac{1}{1+(s/K)^2}, \qquad c_{pm2}(s) = \exp\left(-(s/K)^2\right) \tag{3}$$

where $K > 0$ is the contrast parameter. By such choices of nonlinear functions, PM PDE (2) avoids the over-smoothing property of the heat equation. Good numerical results coupled with the fact that, theoretically, the PM PDE with diffusion functions (3) is ill-posed [3], generated an enormous interest in the mathematical image processing community, see [2] for a review. Moreover, an anisotropic PDE such as (2) can be considered as a gradient descent of a suitable energy functional [4, 5, 6]. The success of the anisotropic diffusion can be attributed to the fact that the PDE can be effectively discretized [7].

Though the PDE based schemes exhibit good denoising behavior, sometimes they can give artifacts such as staircasing or blocky regions. These drawbacks can occur due to various reasons, the primary one is the use of gradients to control diffusion. To avoid this, there have been efforts to use better control mechanisms for inhibiting diffusion in flat regions of the image. These techniques can be classified into three broad categories: (1) Use spatial or time regularization of the gradients [8, 9, 10, 11] (2) Use a separate PDE to get better diffusion coefficients [12, 13, 14, 15] (3) spatially adaptive diffusion coefficients [16, 17, 18, 19, 20, 21, 22]. Though the spatial regularization reduces the effect of noise in gradient computations, it can still give staircasing effects and can have poor localization of edges. In coupled PDE based schemes, apart from the expense of solving another PDE to get the edge map, it can inherit the problems of the original diffusion PDE. Spatially adaptive diffusion coefficient based scheme tries to balance these issues by providing a robust edge map for the diffusion to act upon. Recently, nonlocal diffusion operators were considered in [23, 24, 25] with corresponding wellposedness results. Another approach is to use higher order diffusion models [26].

Here we consider an adaptive scheme which is based on this methodology. Recently, Barcelos et al [27, 28] considered a well-balanced model inspired by the idea of mean curvature motion [29] and Nordstörm's biased PDE [30] approach. In this paper, we generalize such a model and consider weighted anisotropic diffusion schemes which incorporates adaptive information computed from the image at scale t. Moreover, following Smolka [31] a modification of the image fidelity term is also done to improve the denoising capability of the PDE. Following [27], wellposedness of the proposed scheme is proved using the theory of viscosity solutions. Numerical examples in image denoising are given to highlight the proposed model.

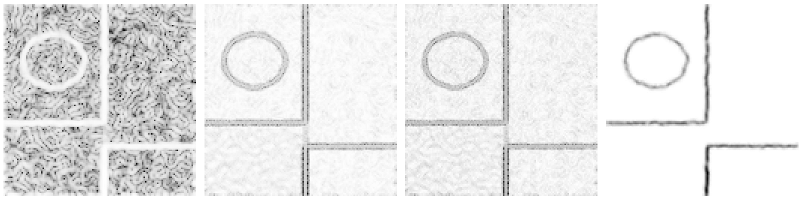


(a) (b) (c)

(d) (e) (f)

Figure 1: Diffusion PDE denoising depends on good edge maps and if the noise persists through the iterations, it leads to staircasing artifacts in the denoised version. (a) Original *Kikis* image used in the experiments (b) Smoothed gradient $|G_\sigma \star \nabla u|$ of the original image, $\sigma = 2$ (Black signifies higher values and white lower) (c) Edge map of the original image computed using the diffusion function c_{pm1} from (3) with $K = 20$ (d) Noisy image obtained by adding Gaussian noise of $\sigma_n = 30$ to the original image (e) Smoothed gradient $|G_\sigma \star \nabla u_0|$ of the noisy image, $\sigma = 2$ (f) Edge map of the noisy image computed using the diffusion function c_{pm1} from (3) with $K = 20$.

The rest of the paper is organized as follows. Section 2.1 introduces the proposed weighted anisotropic diffusion scheme and a modification based on the well-balanced flow model of [27] is presented. Section 4 details the numerical aspects and shows comparison denoising results on noisy images. Finally, Section 5 concludes the paper.



(a) (b) (c) (d)

Figure 2: Edge stopping Vs adaptive diffusion coefficients: (a) Edge stopping function of the noisy image $(1+|G_\sigma \star \nabla u_0|^2)^{-1}$ (b) Inverse gradient based $c(x,|\nabla u|) = \alpha(x)\,|\nabla u|$, where $\alpha(x) = 1/(1+K\,|G_\sigma \star \nabla u_0|^2)$ (c) Slowed diffusion $c(x,|\nabla u|) = (G_\sigma \star \nabla u)/(1+|G_\sigma \star \nabla u|^2/K^2)$ (d) Canny edge detector based $c(x,|\nabla u|) = \alpha(x)\,|\nabla u|$, where $\alpha(x) = 1 - G_\sigma \star Canny(u(x,t))$.

2. Weighted well-balanced anisotropic diffusion

2.1. Well-balanced flow equation

The well-balanced flow (WBF) equation studied by Barcelos et al. [27, 28] is based on total variation and can be generalized to the divergence process such as the Perona-Malik diffusivity:

$$\frac{\partial u}{\partial t} = g\,|\nabla u|\; div\,(c\,(|\nabla u|)\,\nabla u) - \lambda(1-g)(u-u_0) \tag{4}$$

where $g(u \star \nabla G_\sigma) = (1+|G_\sigma \star \nabla u|^2)^{-1}$ is known as the edge stopping function. The pre-smoothing with $G_\sigma(x) = (2\pi\sigma)^{-1}\exp -(|x|^2/2\sigma)$, a Gaussian kernel of width σ, is used to avoid noisy oscillations from the gradient computations. If the diffusion function is $c(s) = s^{-1}$ (total variation (TV) [32]) then we recover the model studied in [27]. This TV diffusion function, in a sense, represents the borderline case from a class of decreasing diffusion functions. More faster decreasing functions can also be used, for example [33], $c(s) = s^{-2}$, though wellposedness results for (4) can not be obtained in these cases.

2.2. Weighted anisotropic diffusion

Figure 1 shows the synthetic image used in our experiments. It consists of homogeneous regions separated by strong edges, gradual slope and a circle object with noisy oscillations. Figures 1(b) & (c) show the smoothed gradient and diffusion function c_{pm1} computed using the original image. These images show that the edge map of the image is captured by the diffusion coefficient and highlights its importance in restoration. The diffusion coefficient c used in the PDE (4) can be influenced greatly by noise and gradient computations can be oscillatory. Figures 1(e)&(f) show the smoothed gradient and diffusion function c_{pm1} values, respectively, computed using the noisy image $|\nabla u_0|$. Clearly, an edge map obtained in this way can lead to diffusion leakage and further iterations can propagate these oscillations which gives staircasing artifacts. Moreover, these gradient based diffusion functions give rise to edge pruning under evolution [34]. Hence, we need to use an adaptive measure which can give a pixel-wise information to the diffusion function $c(x, |\nabla u|)$ in the divergence process. We propose the following class of functions for the diffusion PDE (6):

$$c(x, |\nabla u|) = \alpha(x)\, c_g(|\nabla u|) \tag{5}$$

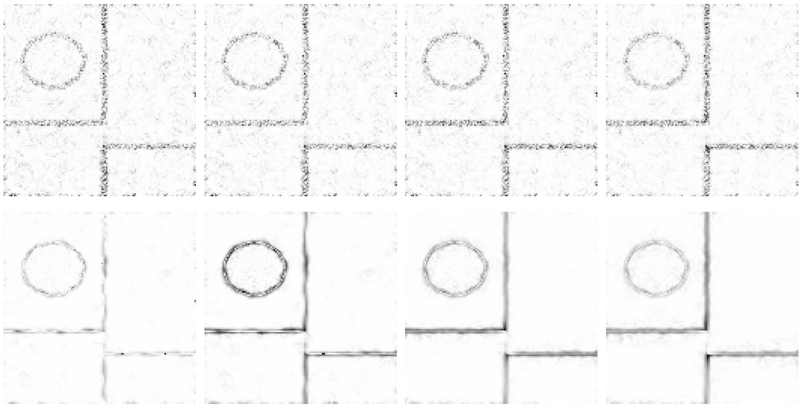
Here, α is the adaptive parameter estimated at each pixel $x \in \Omega$. The function c_g depends on the gradient image $|\nabla u|$ and can be chosen similar to (3). Note that, similar adaptive diffusion function studied in [16] is done for TV gradient function, i.e $c_g(|\nabla u|) = |\nabla u|$. Further, the proposed scheme (6) is modified to include the balance term of [27], and thus provides a well-balanced flow in terms of noise removal and edge preservation. Thus, we consider the following general model (Nordström's biased version [30]) based on PM PDE from Eqn. (2):

$$\frac{\partial u}{\partial t} = g\, div\left(c\left(x, |\nabla u|\right) \nabla u\right) - \lambda(1-g)(u - u_0) \tag{6}$$

where the parameter λ balances the fidelity term and the usual divergence process. Here, we made the diffusion function $c(x, |\nabla u|)$ to depend on the spatial variable $x \in \Omega$ as well as the magnitude of the gradient $|\nabla u|$, which implies the introduction of inhomogeneity into the PDE. For this reason we call the PDE in Eqn (6) as weighted and well-balanced flow (WWBF) equation.

2.3. Choice of diffusion function, weight and other issues

The original Perona-Malik diffusion functions (3) represent two different behaviors with respect to the way the diffusion propagation is carried out. The



(a) $t = 20$ (b) $t = 40$ (c) $t = 100$ (d) $t = 400$

Figure 3: Effect of the balancing (fidelity term) on denoising the $Kikis$ noisy image ($\sigma_n = 30$) using the PM PDE (2) with c_{pm1}. Each image shows the fidelity at different time stamps $t = 20, 40, 100, 200$ (Black signifies higher values and white lower). Top row: classical fidelity $(1 - g) \left|u(x,t) - u_0(x)\right|$ Bottom row: adaptive fidelity $(1 - g) \left|u(x,t) - u(x,t-1)\right|$.

c_{pm1} prefers flat regions over edges and can inhibit higher gradients faster than the c_{pm2} function. To make the presentation simple, throughout the article we use the c_{pm1} diffusion function in all the PDEs. There exists various choices [17, 19, 21, 22] for the weight function α in Eqn. (6). The first choice is to use the classical inverse gradient approach [17], $\alpha(x) = (1 + K \left|\nabla u_0(x)\right|^2)^{-1}$, the other two choices are the slowed diffusion approach [35], and the Canny edge detector based parameter [22], $\alpha(x) = 1 - G_\sigma \star Canny(u(x,t))$. Figure 2 illustrate the usage of adaptive diffusion coefficient against the traditional edge stopping function in front of the divergence term. Note that the edge stopping function g acts as the 'rate' of the diffusion whereas the adaptive coefficient α controls the 'amount' of diffusion. In this sense, both the edge stopping function g and the adaptive parameter α give complementary information for solving the denoising problem. Table 1 provides a succinct comparison of different weight functions from the literature with respect to image restoration. We utilize the inverse gradient function as the weight in the numerical experiments reported here and observed similar results with other adaptive parameter based functions.

Table 1: Comparison of different weight functions for image denoising and restoration. Note that G_σ is a Gaussian kernel, $\mathbf{1}_A$ is the indicator function for a set A, χ_c is a smooth edge indicator function, $Var^2_{\mathcal{N}_x}(u)$ is the local variance of the image function u, for more details we refer to the corresponding references.

Ref.	$\alpha(x)$	Advantages	Disadvantages
[21]	$(1+\lvert G_\sigma \star \nabla u_0(x)\rvert^2)^{-1}$	No staircasing artifacts	Small-scale edges lost
[35]	$\mathbf{1}_{(0.5,1]}(G_\sigma \star \nabla u(x,t))/(1+\lvert G_\sigma \star \nabla u\rvert^2/K^2)$	No diffusion at edges	Noise remain along edges
[19]	$1+M_c\chi_c$, $M_c \gg 0$ constant	Edge indication	Excessive blurring
[22]	$1-G_\sigma \star Canny(u(x,t))$	Retains multi-scale edges	Cannot handle high noise
[36]	$\exp\left(-\Theta(Var^2_{\mathcal{N}_x}(u(x,t)),\theta)/\delta\right)$	Contextual discontinuities	Stippled pattern artifacts
[37]	$\mathbf{1}_I+\mathbf{1}_{I^c}\exp\left(-\left(\lvert G_\sigma \star \nabla u(x)\rvert/K\right)^2\right)$	Handles impulse noise	Cannot handle textures

Note that the fidelity term in Eqn. (6) provides a complementary information using the noisy image $u(x,0)=u_0(x)$. To further increase the denoising capability, we can make the classical image fidelity term $(u(x,t)-u(x,0))$ in Eqn. (6) to be adaptive, i.e., $(u(x,t)-u(x,t-1))$, see Smolka [31]. Figure 3 shows the effect of fidelity on denoising the noisy *Kikis* image (Figure 1(d)) using the PM PDE (2) with diffusion coefficient c_{pm1}. Comparing the adaptive approach (Figure 3, bottom row) with the classical fidelity (Figure 3, top row), we can see that the adaptive process keeps edge details as the iteration increases. This, in turn, will aid the proposed WWBF PDE (6) to smooth the noisy image without destroying the salient edges.

Remark 1. *The balancing term parameter λ can also be made adaptive, see Gilboa et al [38]. A spatially adaptive balance parameter $\lambda(x)$ can keep the textural component in the restored image u, while keeping the fidelity constraint.*

Remark 2. *Further generalizations of the well-balanced flow are also possible. For example, the diffusion coefficient can also be made to depend on the image u, i.e., $c(x,u,|\nabla u|)$. Such a generalization can lead to different diffusion flows and can be designed to influence the restoration process.*

The wellposedness of the proposed PDE (6) can be proved using the viscosity solution theory and its discretized version satisfies the usual scale-space properties as well.

3. Theoretical considerations

3.1. Preliminaries

Following [27, 39], we study the proposed PDE

$$\frac{\partial u}{\partial t} = g(G * \nabla u)\, div\,(c\,(x, |\nabla u|)\,\nabla u) - \lambda(1 - g(G * \nabla u))(u - u_0) \tag{7}$$

using the viscosity solution theory of P. L. Lions et al [40]. Here we admit generic convolution kernels G, which, in particular, can be the Gaussian kernels G_σ, and arbitrary spatial dimension $n > 1$.

Throughout this section, we employ Einstein's summation convention. Let us first introduce two auxiliary functions depending on x and p from $\mathbb{R}^n$, a symmetric-matrix-valued one a and a vector one χ. We denote

$$a_{ij}(x, p) = c(x, |p|)\delta_{ij} + c_y(x, |p|)\frac{p_i p_j}{|p|}, \tag{8}$$

$$\chi_i(x, p) = \frac{\partial c(x, |p|)}{\partial x_i}. \tag{9}$$

Here δ_{ij} is Kronecker's delta, and c_y is the partial derivative of $c(x, y)$ with respect to the second variable.

As usual, for the sake of simplification of the presentation, we consider the case of spatially periodic boundary conditions [39] for Eqn. (7). Namely, we assume that there is an orthogonal basis $\{b_i\}$ in $\mathbb{R}^n$ so that

$$u(\cdot, x) = u(\cdot, x + b_i),\ x \in \mathbb{R}^n,\ i = 1, \ldots, n. \tag{10}$$

The problem is complemented with the initial condition

$$u(0, x) = u_0(x), \tag{11}$$

where $x \in \mathbb{R}^n$, and u_0 is Lipschitz and satisfies (10). Of course, c (and thus a and χ) should also satisfy the same spatial periodicity restriction (with respect to x but not to y or p).

Let us introduce the following algebraic notion. Given a diagonal matrix B, let $mod(B)$ be the matrix whose entries are the absolute values of the entries of B. Furthermore, if B is an arbitrary symmetric matrix, it can be represented as $Q^\top DQ$, where D is a diagonal matrix and Q is an orthogonal one. Then we define $mod(B) = Q^\top mod(D)Q$. It is straightforward to check that this definition does not depend on a particular choice of D and Q. Observe also that $mod(B)$ is always positive-semidefinite, whereas $mod(B) = B$ when B itself is positive-semidefinite.

We make the following assumptions:

$$a, \chi \text{ are continuous, bounded, periodic in } x \text{, continuously differentiable in } x, \tag{12}$$

$$\text{and their } x\text{-derivatives are uniformly (w.r.t. } p\text{) bounded,} \tag{13}$$

$$a_{ij}(x,p)\xi_i\xi_j \geq C\left[mod\left(\frac{\partial a(x,p)}{\partial x_k}\right)\right]_{ij}\xi_i\xi_j,\ k = 1,\dots,n,\ \xi, x, p \in \mathbb{R}^n, \tag{14}$$

$$g : \mathbb{R}^n \to \mathbb{R},\ 0 \leq g \leq 1,\ \sqrt{g} \text{ is Lipschitz,} \tag{15}$$

$$G \in W_1^2(\mathbb{R}^n) \text{ (note that we do not assume it to be space-periodic),} \tag{16}$$

$$\lambda \geq 0. \tag{17}$$

Here and below C stands for a generic positive constant, which can take different values in different lines.

Definition 1 (Viscosity solution)**.** *A function u from the space*

$$C([0,T]\times\mathbb{R}^n) \cap L_\infty(0,T,W_\infty^1(\mathbb{R}^n)) \tag{18}$$

is a viscosity sub-/supersolution to (7), (10), (11) if, for any $\phi \in C^2([0,T]\times\mathbb{R}^n)$ and any point $(t_0,x_0) \in (0,T]\times R^n$ of local maximum/minimum of the function

$u - \phi$, one has

$$\begin{aligned}\frac{\partial\phi(t_0,x_0)}{\partial t} &- g((u*\nabla G)(t_0,x_0))\,div\,(c\,(x_0,|\nabla\phi(t_0,x_0)|)\,\nabla\phi(t_0,x_0)) \\ &+ \lambda(1-g((u*\nabla G)(t_0,x_0)))(u(t_0,x_0)-u_0(x_0)) \le 0 \;/\; \ge 0, \end{aligned} \tag{19}$$

and equalities (10), (11) hold in the classical sense. A viscosity solution is a function which is both a subsolution and a supersolution.

3.2. Main result

Theorem 1. *i) The problem (7), (10), (11) has a viscosity solution in class (18) for every positive T. Moreover,*

$$\inf_{\mathbb{R}^n} u_0 \le u(t,x) \le \sup_{\mathbb{R}^n} u_0. \tag{20}$$

ii) Assume that

$$\left|\left(\sqrt{a(x,p)} - \sqrt{a(z,p)}\right)_{ij}\right| \le C|x-z|, \;\; x,z,p \in \mathbb{R}^n. \tag{21}$$

Here $\sqrt{\ }$ is the square root of a positive-semidefinite symmetric matrix [41]. Then the solution is unique. Moreover, for any two viscosity solutions u and v to (7), the following estimate holds

$$\sup_{\mathbb{R}^n} |u(t,\cdot) - v(t,\cdot)| \le \Phi(t) \sup_{\mathbb{R}^n} |u(0,\cdot) - v(0,\cdot)| \tag{22}$$

with some non-decreasing continuous scalar function Φ dependent on u and v.

Proof. Note that (20) is a direct consequence of the definition of viscosity solution: to get the second inequality, one can put $\phi(t,x) = \delta t$, then, at the point (t_0,x_0), $t_0 > 0$, of the global maximum of $u(t,x) - \delta t$, (19) gives $\delta + \lambda(1-g((u*\nabla G)(t_0,x_0)))(u(t_0,x_0)-u_0(x_0)) \le 0$, whence $u(t_0,x_0) < u_0(x_0)$, so we get a contradiction since $u(t_0,x_0) - \delta t_0 \ge u_0(x_0)$ due to the fact that (t_0,x_0) is the global maximum point of $u(t,x) - \delta t$; thus the function $u(t,x) - \delta t$ attains its global maximum at $t = 0$, and it remains to let $\delta \to +0$; similarly one derives the first one.

Now, we establish a *formal a priori estimate* for $\sup_{\mathbb{R}^n} |\nabla u|$. Observe that (7) is equivalent to

$$\begin{aligned}\frac{\partial u}{\partial t} = g(u * \nabla G) \left[a_{ij}(x, \nabla u) u_{x_i x_j} + \chi_i(x, \nabla u) u_{x_i}\right]& \\ - \lambda(1 - g(u * \nabla G))(u - u_0).& \qquad (23)\end{aligned}$$

Fix T. Differentiating (23) with respect to each x_k, $k = 1, \dots, n$, multiplying by $2u_{x_k}$, and adding the results, we get

$$\begin{aligned}\mathcal{L}(|\nabla u|^2) := \frac{\partial |\nabla u|^2}{\partial t} &- g a_{ij}(x, \nabla u) \frac{\partial^2}{\partial x_i \partial x_j} |\nabla u|^2 \\ &- g \frac{\partial a_{ij}(x, \nabla u)}{\partial p_l} u_{x_i x_j} \frac{\partial}{\partial x_l} |\nabla u|^2 \\ &- g \chi_i(x, \nabla u) \frac{\partial}{\partial x_i} |\nabla u|^2 - g \frac{\partial \chi_i(x, \nabla u)}{\partial p_l} u_{x_i} \frac{\partial}{\partial x_l} |\nabla u|^2 \\ = -2g a_{ij}(x, \nabla u) u_{x_k x_i} u_{x_k x_j} &+ 2 \nabla g(u * \nabla G) \cdot \left(u * \frac{\partial \nabla G}{\partial x_k}\right) a_{ij}(x, \nabla u) u_{x_i x_j} u_{x_k} \\ + 2g \frac{\partial a_{ij}(x, \nabla u)}{\partial x_k} u_{x_i x_j} u_{x_k} &+ 2 \nabla g(u * \nabla G) \cdot \left(u * \frac{\partial \nabla G}{\partial x_k}\right) \chi_i(x, \nabla u) u_{x_i} u_{x_k} \\ + 2g \frac{\partial \chi_i(x, \nabla u)}{\partial x_k} u_{x_i} u_{x_k} &- 2\lambda(1 - g) u_{x_k} u_{x_k} + 2\lambda(1 - g)(u_0)_{x_k} u_{x_k} \\ &+ 2\lambda \nabla g(u * \nabla G) \cdot \left(u * \frac{\partial \nabla G}{\partial x_k}\right)(u - u_0) u_{x_k}. \qquad (24)\end{aligned}$$

At this point, we need the following generalization of [39, Lemma 2.6].

Lemma 1. *Let A and B be quadratic matrices of order n. Assume that B is symmetric, and there is a constant $M \geq 0$ such that*

$$M A_{ij} \xi_i \xi_j \geq mod(B)_{ij} \xi_i \xi_j, \ \forall \xi \in \mathbb{R}^n. \qquad (25)$$

Then for any matrix U (of the same order but not necessarily symmetric) one has

$$Tr^2(BU^\top) \leq M \|B\| Tr(UAU^\top), \qquad (26)$$

where $\|\cdot\|$ denotes the operator norm of a matrix.

Proof. Formulas (25) and (26) are invariant with respect to orthogonal changes of bases. Thus, without loss of generality we may assume that B is already diagonalized by an orthogonal transform. Then

$$Tr^2(BU^\top) = (B_{ii}U_{ii})^2 \le \|B\||B_{ii}|U_{ii}^2$$

$$= \|B\|(mod(B))_{ii}U_{ii}^2 \le \|B\|(mod(B))_{ii}U_{ki}U_{ki}$$

$$= \|B\|(mod(B))_{ij}U_{ki}U_{kj} \le M\|B\|A_{ij}U_{ki}U_{kj} = M\|B\|Tr(UAU^\top).$$

□

This lemma gives opportunity to discharge the undesired influence of the second and the third terms in the right-hand side of (24). For the third one, due to the lemma, (14) and Cauchy's inequality, we have

$$\left|2g\frac{\partial a_{ij}(x,\nabla u)}{\partial x_k}u_{x_ix_j}u_{x_k}\right| \le Cg\left|u_{x_k}\right|\sqrt{a_{ij}(x,\nabla u)u_{x_kx_i}u_{x_kx_j}}$$
$$\le ga_{ij}(x,\nabla u)u_{x_kx_i}u_{x_kx_j} + C|\nabla u|^2. \quad (27)$$

Since our assumptions yield

$$\left|u * \frac{\partial \nabla G}{\partial x_k}\right| \le C, \quad (28)$$

and

$$|\nabla g| \le C\sqrt{g}, \quad (29)$$

an application of the lemma with $A = B = a$ and $M = 1$ implies

$$\left|2\nabla g(u*\nabla G)\cdot\left(u*\frac{\partial\nabla G}{\partial x_k}\right)a_{ij}(x,\nabla u)u_{x_ix_j}u_{x_k}\right|$$
$$\le C\left|u_{x_k}\right|\sqrt{ga_{ij}(x,\nabla u)u_{x_kx_i}u_{x_kx_j}}$$
$$\le ga_{ij}(x,\nabla u)u_{x_kx_i}u_{x_kx_j} + C|\nabla u|^2. \quad (30)$$

The sum of the absolute values of the subsequent terms of the right-hand side of (24) does not exceed $C(1+|\nabla u|^2)$. Thus,

$$\mathcal{L}(|\nabla u|^2) \le C(1+|\nabla u|^2), \quad (31)$$

so

$$\mathcal{L}\left(e^{-Ct}(1+|\nabla u|^2)\right) \le 0. \quad (32)$$

From the weak maximum principle for the weakly parabolic operator $\mathcal{L}$ one easily concludes that

$$|\nabla u|^2 \leq C. \tag{33}$$

Using (20) and (33), by means of the approach from [39] we can get the uniform Hölder estimate

$$|u(t,x) - u(s,x)|^2 \leq C|t-s|. \tag{34}$$

Then, following [43, 39], we approximate our problem by well-posed ones in the sense of [42, Chapter 5]. Due to (20), (33) and (34), the solutions of these problems are uniformly bounded and equicontinuous on $[0,T] \times \mathbb{R}^n$. Then we can select a uniformly converging sequence of approximate solutions, and pass to the limit in the viscosity sense using the general consistency/stability results from [40]. The uniqueness of solutions follows from the stability estimate (22). This bound may be shown by revisiting the proof of a similar bound in [43, 39]. We only point out that the matrix Γ [39, p. 159] is replaced by

$$\Gamma_* = \begin{pmatrix} g_1\Lambda_1 & \sqrt{g_1 g_2}\sqrt{\Lambda_1}\sqrt{\Lambda_2} \\ \sqrt{g_1 g_2}\sqrt{\Lambda_2}\sqrt{\Lambda_1} & g_2\Lambda_2 \end{pmatrix}, \tag{35}$$

where

$$\Lambda_1 = a\left(x_0, \frac{|x_0 - y_0|^2(x_0 - y_0)}{\varepsilon}\right), \ \Lambda_2 = a\left(y_0, \frac{|x_0 - y_0|^2(x_0 - y_0)}{\varepsilon}\right),$$

and the notation within is taken from [39]. Note that the $2n \times 2n$-matrix Γ_* is symmetric and positive-semidefinite.

□

4. Numerical Results

4.1. Comparison with other schemes

The proposed scheme is compared with related diffusion based denoising schemes from the literature. To make a fair comparison we utilize the same diffusion function c_{pm1} from (3) in all the compared schemes and the contrast parameter K is fixed using the original criteria given in [1], see [44, 45] for other choices. Moreover, the edge stopping function $g(\xi) = (1 + |\xi|)^{-1}$ is fixed wherever applicable and the classical fidelity term is utilized unless otherwise stated.

(a) Perona and Malik [1] - Anisotropic Diffusion (AD) Eqn. (2) with c_{pm1} in (3):

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{1 + |\nabla u|^2 / K^2}\right)$$

(b) Catté et al [8] - Smoothed Gradient based anisotropic diffusion (SG) with c_{pm1} in (3):

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{1 + |\nabla G_\sigma \star u|^2 / K^2}\right)$$

(c) Rudin et al [32] - Total Variation (TV) (2) with $c(s) = (\epsilon + s^2)^{-1/2}$, $\epsilon = 10^{-6}$:

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla u}{\sqrt{\epsilon + |\nabla u|^2}}\right)$$

(d) El Falah and Ford [29] - Mean Curvature Motion (MCM), Eqn. (4) with $\lambda = 0$:

$$\frac{\partial u}{\partial t} = \frac{1}{1 + |\nabla u|^2}\, div\left(\frac{\nabla u}{1 + |\nabla u|^2 / K^2}\right)$$

(e) Barcelos et al [27] - Well-Balanced Flow (WBF):

$$\frac{\partial u}{\partial t} = g(|\nabla G_\sigma \star u|)\, div\left(\frac{\nabla u}{1 + |\nabla u|^2 / K^2}\right) - \lambda(1 - g(|\nabla G_\sigma \star u|))(u - u_0)$$

(f) Shi and Chang [9] - Modified Smoothed Gradient based anisotropic diffusion (MSG):

$$\frac{\partial u}{\partial t} = |\nabla G_\sigma \star u|\; div\left(\frac{\nabla G_\sigma \star u}{|\nabla G_\sigma \star u|}\right) - |\nabla G_\sigma \star u|\, \lambda(G_\sigma \star u - u_0)$$

Further, similar adaptive schemes which utilize different diffusion coefficient functions are also compared.

(a) Weickert [46] - Edge Enhancing Diffusion (EED):

PM PDE (2) with the diffusion function:

$$c(|\nabla u|) = \begin{cases} \exp\left(-0.234\ |\nabla u|\right) & \text{if } |\nabla u| \geq K \\ 0 & \text{if } |\nabla u| < K \end{cases}$$

(b) Weickert [47] - Coherence Enhancing Diffusion (CED):

PM PDE (2) with the diffusion function constructed using the structure tensor, see [47] for more details. The eigenvalues of D are chosen as, for μ_1, μ_2 eigenvalues of the structure tensor, $\alpha \in (0,1)$, $C > 0$: $\lambda_1 = \alpha$, and

$$\lambda_2 = \begin{cases} \alpha & \text{if } \ \mu_1 = \mu_2 \\ \alpha + (1-\alpha)\exp\left(\frac{-C}{(\mu_1-\mu_2)^2}\right) & \text{else} \end{cases}$$

(c) Kačur and Mikula [48, 35] - Slowed Anisotropic Diffusion (SAD):

$$\frac{\partial u}{\partial t} = div\left(\frac{\nabla G_\sigma \star u}{1 + |\nabla G_\sigma \star u|^2 / K^2} \nabla \beta(x,u)\right)$$

with $\beta(x,u) = 0$ for $u \in [0, 0.5]$ and $\beta(x,u) = u$ for $u \in (0.5, 1]$.

(d) Strong [16] - Adaptive TV (ATV):

$$\frac{\partial u}{\partial t} = div\left(\frac{\alpha(x)\nabla u}{\sqrt{\epsilon + |\nabla u|^2}}\right)$$

with $\alpha(x) = (1 + |\nabla u_0|)^{-1}$, $\epsilon = 10^{-6}$.

(e) Kusnezow et al [19] - Adaptive Linear Diffusion (ALD):

$$\frac{\partial u}{\partial t} = \alpha\, div\left(\nabla u\right)$$

with $\alpha = (1 + M_c \chi_c)$, $M_c \gg 0$ constant and χ_c is a smooth edge indicator function.

(f) Prasath and Singh [22] - Edge detector based Anisotropic Diffusion (EAD)

$$\frac{\partial u}{\partial t} = div\left(\frac{\alpha(x)\nabla u}{1 + |\nabla u|^2 / K^2}\right)$$

with $\alpha(x) = 1 - G_\sigma \star C(u(x,t))$, C - Canny edge detector output.

In the comparison results, apart from using the proposed adaptive fidelity term based WWBF (see Eqn. (6) and Section 2.3),

$$\frac{\partial u(x,t)}{\partial t} = g\, div\left(c\left(x, |\nabla u|\right) \nabla u(x,t)\right) - \lambda(1-g)(u(x,t) - u(x,t-1)),$$

we also utilize the weighted Linear Diffusion (WLD) - using the proposed weight in a linear diffusion framework,

$$\frac{\partial u}{\partial t} = g\, div\left(\alpha(x)\nabla u\right) - \lambda\left(1-g\right)\left(u(x,t) - u(x,t-1)\right).$$

4.2. Implementation details

The additive operator splitting (AOS) scheme which is proven to be effective in diffusion PDE based image processing [7] is used to implement the schemes. The images were scaled to the interval $[0,1]$. It can be described briefly as follows: In 1-D with matrix-vector notation, the iterative scheme is,

$$U^{t+1} = \left[1 - \tau A(U^t)\right]^{-1} U^t,$$

where τ is the time step, $A(U^t) = [a_{ij}(U^t)]$, and

$$a_{ij}(U^t) := \begin{cases} \frac{\gamma_i^t + \gamma_j^t}{2h^2} & j \in \mathcal{N}_i \\ -\sum_{k \in \mathcal{N}_i} \frac{\gamma_i^t + \gamma_k^t}{2h^2} & j = i \\ 0 & \text{otherwise} \end{cases}$$

with $\gamma_i = \alpha_i g_i$ and h discretization step size. For n-D images the semi-implicit scheme is written as

$$U^{t+1} = \left[1 - \tau \sum_{l=1}^{n} A_l(U^t)\right]^{-1} U^t. \tag{36}$$

The matrix $A_l = (a_{ijl})_{ij}$ corresponds to derivatives along the l-th coordinate axis.

Remark 3. *The spatial step size $h = 1$ is fixed as the pixel grid has the natural spacing of size one. Further the time step $\tau = 0.2$, pre-smoothing parameter $\sigma = 1$, and fidelity parameter $\lambda = 1$ are fixed for all the experiments reported here.*

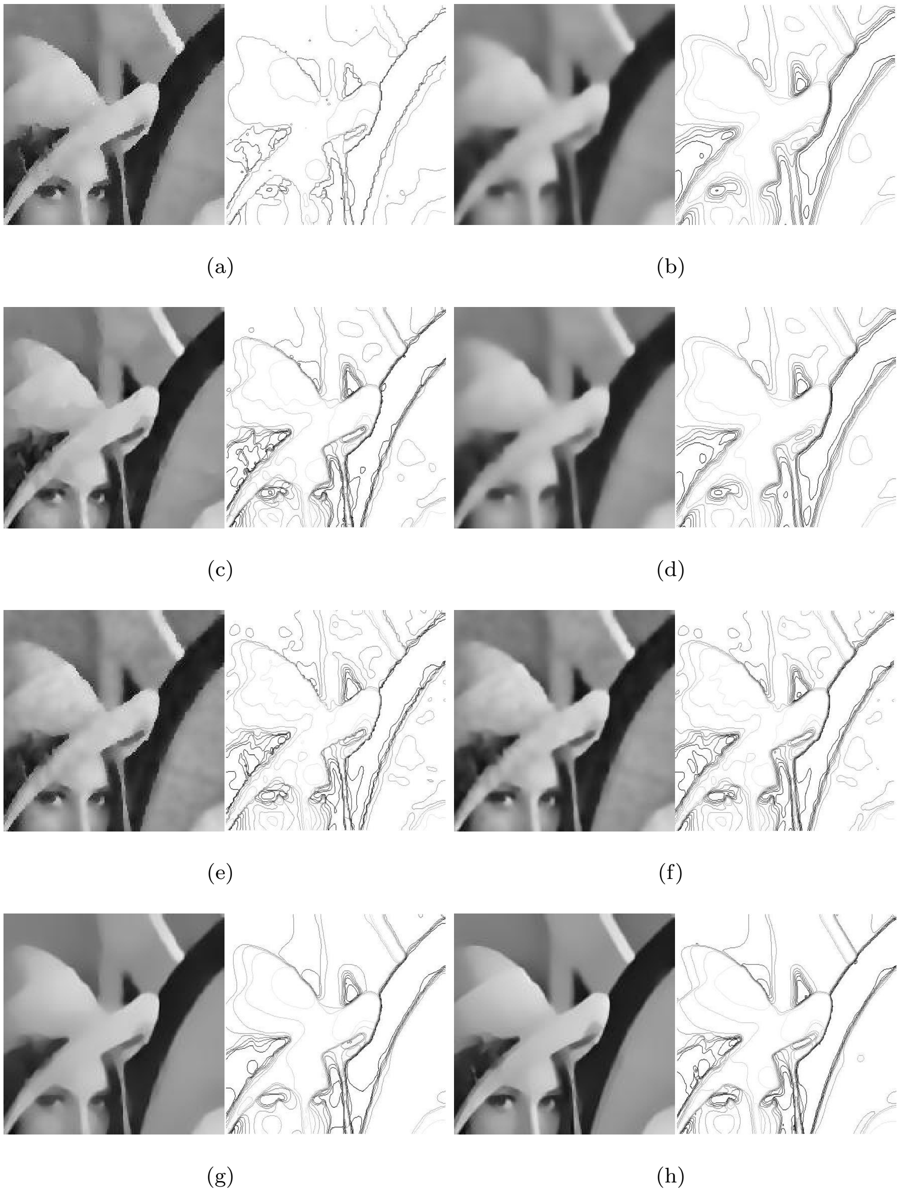
Remark 4. *Under the AOS type discretization* (36)*, the proposed WWBF scheme* (6) *satisfies the usual scale space properties, see [7] for more details. Moreover the maximum-minimum principle also holds, see Theorem 1.*

4.3. Visual comparison

Figure 4 shows the comparison of non-adaptive diffusion schemes based restoration results for a noisy (Gaussian noise, $\sigma_n = 25$) *Lena* gray-scale image. In each pair, left image shows the 156×156 crop of the restored image and the right image shows the contour view to highlight the movement of level-sets under different schemes. Note that, the proposed approach gives better result even with linear diffusion, see Figure 4(g) which corresponds to WLD scheme result. As can be seen by comparing the contour maps of each scheme, the proposed scheme's result in Figure 4(h) gives better result in terms denoising as well as staircasing artifact free restoration.

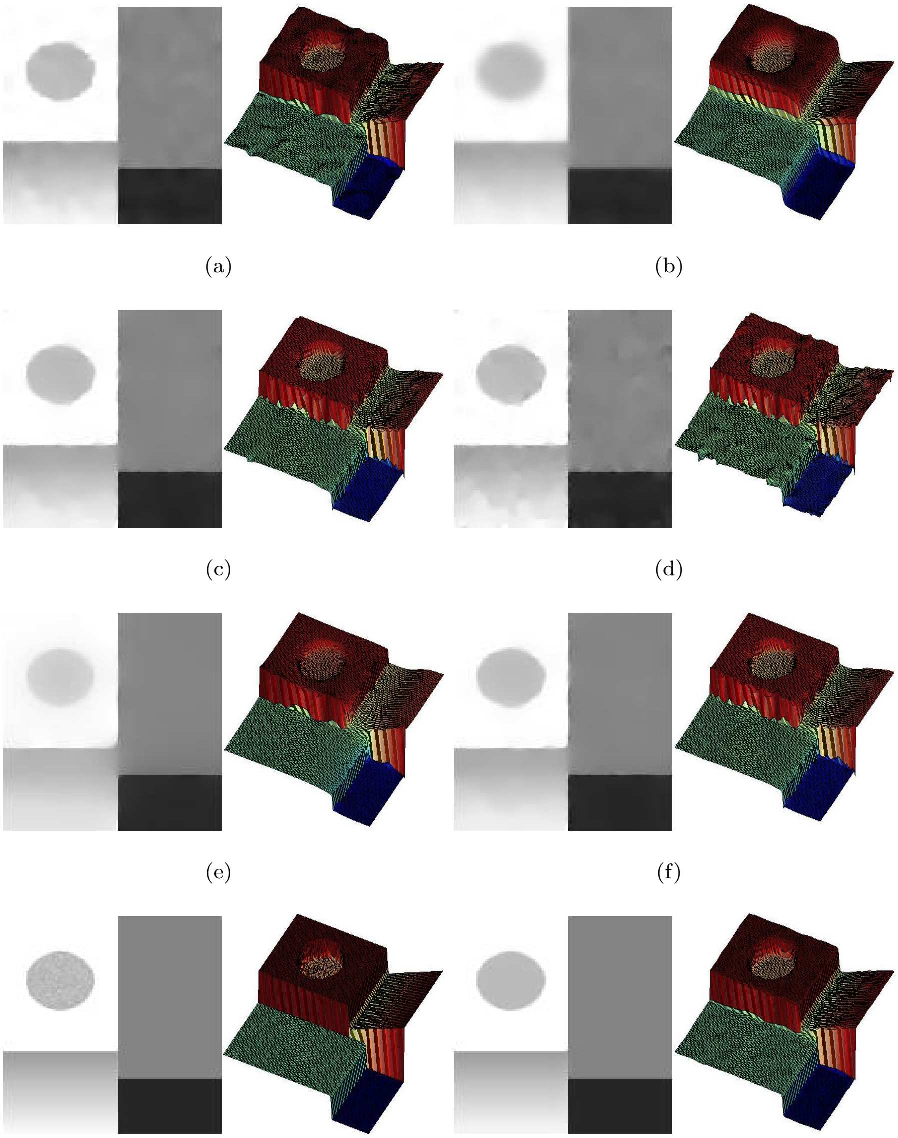
To compare the adaptive diffusion schemes in a fair manner we utilize a test image synthetically generated consist of a slope, strong edges and a circle with oscillations. Figure 5 shows the comparison results for the noisy $Kikis$ image ($\sigma_n = 30$ is added to the original image, see Figure 1(d)) by different adaptive diffusion schemes. The Perona-Malik, TV based schemes such as EED, CED, SAD, ATV inherit the original staircasing artifacts whereas WWBF performs better than other schemes in terms of edge preservation without oscillations, see Figure 5(h).

Finally, to show the effect of the adaptive fidelity term in different adaptive schemes we perform experiments on a synthetic $Circles$ gray-scale image which has multiple circular regions with different piecewise constant regions. Figure 6 shows the comparison of the adaptive schemes SAD, ATV, ALD, and EAD with the same adaptive fidelity chosen as in our WWBF scheme, i.e., $(u(x,t) - u(x,t-1))$. As can be seen the WWBF scheme preserves edges without any blocky artifacts. Moreover, the adaptive fidelity term captures the circular edges thereby balances the adaptive diffusion near the edges. Supplementary MATLAB .fig files are provided to show 3D visualizations of resultant images



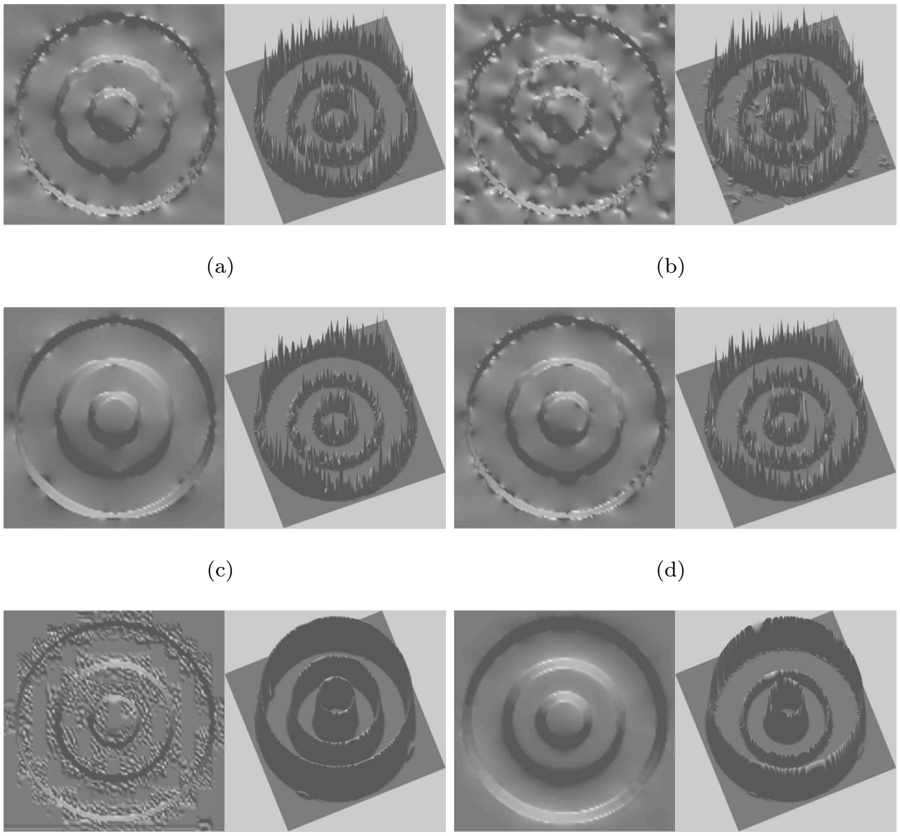
(a) (b) (c) (d) (e) (f) (g) (h)

Figure 4: Comparison results for *Lena* image, cropped 156 × 156 image (in each sub-figure, the right image shows the contour view of the left image). (a) AD [1] (b) SG [8] (c) TV [32] (d) MCM [29] (e) WBF [27] (f) MSG [9] (g) Proposed scheme with linear diffusion (WLD) (h) Proposed scheme with nonlinear diffusion (WWBF).



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 5: Adaptive schemes comparison results on *Kikis* 128×128 synthetic image, (in each sub-figure, the right image shows the surface form of the left image): (a) EED [46] (b) CED [47] (c) SAD [48] (d) ATV [16] (e) ALD [19] (f) EAD [22] (g) Original image and its surface form given for comparison (h) Proposed scheme with nonlinear diffusion (WWBF).



(a) (b)

(c) (d)

(e) (f)

Figure 6: Adaptive schemes comparison results on *Circles* 120×120 synthetic image, (in each sub-figure, right image shows the resultant and left image adaptive fidelity term at the final iteration in surface format): (a) SAD [48] (b) ATV [16] (c) ALD [19] (d) EAD [22] (e) Original image and its edge map given in surface format for comparison. Note that artifacts are due to jpeg compression which appear near edges. (f) Proposed scheme with nonlinear diffusion (WWBF). Supplementary MATLAB .fig files are provided to show 3D visualizations of resultant images shown here.

shown on the left of each sub-figure.

Remark 5. *Other non-adaptive diffusion schemes such as AD, SG, TV, MCM, WBF, MSG and directional diffusion models such as EED, CED do not utilize an adaptive weight as in our case (see Eqn. (5)). Moreover, the adaptive data fidelity term did not provide any visually improved denoising results for these schemes, hence we omit the images in Figure 6 for brevity.*

4.4. Quantitative comparison and discussion

To compare the schemes quantitatively we utilize two commonly used error metrics in the image denoising literature, one is the classical peak signal to noise ratio (PSNR) [2], and the other is the mean structural similarity measure (MSSIM) [49]:

1. PSNR is given in decibels (dB). A difference of $0.5\,dB$ can be identified visually. Higher PSNR value indicates optimum denoising capability.

   $$\text{PSNR}(u) := 20 * \log 10\left(\frac{u_{max}}{\sqrt{MSE}}\right) dB$$

   where $\text{MSE} = (mn)^{-1}\sum\sum(u - u_0)$, $m \times n$ denotes the image size, u_{max} denotes the maximum value, for example in 8-bit images $u_{max} = 255$.
2. MSSIM index is in the range $[0, 1]$. The MSSIM value near one implies the optimal denoising capability of the scheme [49] and is mean value of the SSIM metric. The SSIM is calculated between two windows ω_1 and ω_2 of common size $N \times N$,

   $$\text{SSIM}(\omega_1, \omega_2) = \frac{(2\mu_{\omega_1}\mu_{\omega_2} + c_1)(2\sigma_{\omega_1\omega_2} + c_2)}{(\mu^2_{\omega_1} + \mu^2_{\omega_2} + c_1)(\sigma^2_{\omega_1} + \sigma^2_{\omega_2} + c_2)}$$

   where μ_{ω_i} the average of ω_i, $\sigma^2_{\omega_i}$ the variance of ω_i, $\sigma_{\omega_1\omega_2}$ the covariance, c_1, c_2 stabilization parameters, see [49] for more details.

Table 2 shows the comparison results using these two metrics for all schemes without data adaptive fidelity term. Corresponding PSNR and MSSIM values are given for each of the schemes and clearly our scheme performs better than or on par with other diffusion schemes in general. We also include comparison results with corresponding data adaptive fidelity term described in Section 2.3. As can be noted, the proposed scheme performs well for a variety of images (taken from the standard test images USC-SIPI database) for both data fidelity versions. Note that the PSNR values are closer together when adaptive fidelity is used (SAD, ATV, ALD, EAD, and our WWBF) in Table 2, but MSSIM values indicate a better performance of the proposed approach. Thus, the proposed adaptive WWBF flow preserves salient structures (edges) when compared with other nonlinear heat diffusion flows. The *Baboon* image consist of texture parts and hence the proposed WWBF scheme can not obtain optimal PSNR/MSSIM values. To alleviate this a spatially adaptive fidelity parameter $\lambda = \lambda(x)$ can be incorporated, see Section 2.3. Following [28] automatic selection of parameters is one of the current research being carried out. Moreover, the image restoration model studied here can be used in other image processing algorithms such as inpainting [50, 51] and edge detection [27] as well.

5. Conclusions

Well-balanced flow is based on a nonlinear diffusion PDE which is utilized in image noise removal and edge detection successfully. In this paper, a new variant of the flow is considered by using weights in the divergence diffusion process. This improves the denoising capabilities as well as the multi-scale detail preservation of the corresponding PDE. Numerical experiments on noisy images shows the proposed scheme's performs well on a variety of images. Extensive experiments indicate the improvements over other classical diffusion and adaptive diffusion schemes.

Acknowledgement

The authors would like to thank the anonymous referees for their constructive comments which improved the paper considerably. Part of this work was done while the first author was visiting Institute for Pure and Applied Mathematics (IPAM), University of California Los Angeles (UCLA), USA. The first

Table 2: PSNR (dB) and MSSIM comparison for standard test images with and without adaptive fidelity term for different diffusion based schemes. Noisy image is obtained by adding Gaussian noise of strength $\sigma_n = 25$ to the original image of size 256×256 except for the image *Kikis* which has $\sigma_n = 30$ and size 128×128. Each row indicates the PSNR/MSSIM values for different test images. Overline indicate the PDE is used with adaptive data-fidelity and best results are indicated by boldface.

Scheme	Ref.	*Kikis*	*Lena*	*House*	*Peppers*	*Baboon*
Noisy		18.56/0.1683	20.14/0.3866	20.14/0.2732	20.14/0.3426	20.14/0.4643
AD	[1]	33.32/0.9081	26.32/0.7752	28.87/0.8300	27.37/0.8170	23.48/0.4687
SG	[8]	29.09/0.9036	23.26/0.6708	24.94/0.7657	23.09/0.7291	22.35/0.3653
TV	[32]	33.47/0.9414	27.05/0.7951	30.18/0.8520	28.30/0.8389	23.61/0.4899
MCM	[29]	30.87/0.9238	23.97/0.6943	25.89/0.7855	24.02/0.7501	22.44/0.3723
WBF	[27]	33.19/0.9111	26.46/0.7827	28.94/0.8286	27.63/0.8289	23.54/0.4802
MSG	[9]	33.23/0.9273	26.53/0.7826	29.30/0.8370	27.31/0.8327	23.34/0.4627
EED	[46]	35.23/0.9530	27.23/0.7980	30.68/0.8554	28.44/**0.8435**	23.47/0.4685
CED	[47]	30.87/0.9238	23.97/0.6943	25.89/0.7855	24.02/0.7501	22.44/0.3723
SAD	[48]	34.57/0.9593	25.85/0.7559	29.18/0.8375	26.93/0.8030	22.90/0.4133
ATV	[16]	33.68/0.9435	27.26/0.7972	30.39/0.8541	28.51/0.8410	**23.82**/0.4920
ALD	[19]	28.48/0.9378	20.70/0.5965	23.23/0.7346	20.58/0.6248	20.84/0.3105
EAD	[22]	34.24/0.9600	24.88/0.7247	28.17/0.8221	25.72/0.7719	22.47/0.3762
WLD		32.81/0.9423	23.87/0.7126	27.03/0.8088	23.91/0.7306	20.73/0.3557
WWBF		37.00/0.9499	27.12/0.7815	30.92/0.8584	28.27/0.8109	22.98/0.4417
$\overline{\text{SAD}}$		34.45/0.9601	24.05/0.7595	27.82/0.8409	24.36/0.8109	20.00/0.4211
$\overline{\text{ATV}}$		30.96/0.9481	24.69/**0.8028**	28.93/0.8581	26.09/0.8483	22.67/**0.5004**
$\overline{\text{ALD}}$		26.09/0.9407	19.58/0.6053	21.36/0.7562	18.84/0.6490	19.08/0.3279
$\overline{\text{EAD}}$		33.78/0.9658	22.80/0.7203	26.77/0.8339	24.96/0.7740	21.58/0.3836
$\overline{\text{WLD}}$		33.71/0.9652	24.90/0.7269	28.63/0.8293	25.77/0.7692	22.40/0.3693
$\overline{\text{WWBF}}$		**38.54/0.9696**	**27.42**/0.7965	**31.27/0.8621**	**28.73**/0.8356	23.46/0.4533

author thanks the IPAM institute for their great hospitality and support during the visit. The second author was supported by CMUC (University of Coimbra) and FCT (Portugal), and through European program COMPETE/FEDER.

References

[1] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7) (1990) 629–639. doi:10.1109/34.56205.

[2] G. Aubert, P. Kornprobst, Mathematical problems in image processing: Partial differential equation and calculus of variations, Springer-Verlag, New York, USA, 2006.

[3] S. Kichenassamy, The Perona-Malik paradox, SIAM Journal on Applied Mathematics 57 (5) (1997) 1328–1342. doi:10.1137/S003613999529558X.

[4] Y.-L. You, W. Xu, A. Tannenbaum, M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing, IEEE Transactions on Image Processing 5 (11) (1996) 1539–1553. doi:10.1109/83.541424.

[5] C. A. Z. Barcelos, Y. Chen, Heat flows and related minimization problem in image restoration, Computers & Mathematics with Applications 39 (5–6) (2000) 81–97. doi:10.1016/S0898-1221(00)00048-1.

[6] T. Barbu, V. Barbu, V. Biga, D. Coca, A PDE variational approach to image denoising and restoration, Nonlinear Analysis: Real World Applications 10 (3) (2009) 1351–1361. doi:10.1016/j.nonrwa.2008.01.017.

[7] J. Weickert, B. M. H. Romeny, M. A. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, IEEE Transactions on Image Processing 7 (3) (1998) 398–410. doi:10.1109/83.661190.

[8] V. Catte, P. L. Lions, J.-M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, SIAM Journal on Numerical Analysis 29 (1) (1992) 182–193. doi:10.1137/0729012.

[9] Y. Shi, Q. Chang, New time dependent model for image restoration, Applied Mathematics and Computation 179 (1) (2006) 121–134. doi:10.1016/j.amc.2005.11.085.

[10] H. Amann, Time-delayed Perona-Malik type problems, Acta Mathematica Universitatis Comenianae LXXVI (1) (2007) 15–38.

[11] E. Tadmor, P. Athavale, Multiscale image representation using novel integro-differential equations, Inverse Problems and Imaging 3 (4) (2009) 693–710. doi:10.3934/ipi.2009.3.693.

[12] M. Nitzberg, T. Shiota, Nonlinear image filtering with edge and corner enhancement, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (8) (1992) 826–833. doi:10.1109/34.149593.

[13] Y. Chen, C. A. Z. Barcelos, B. A. Mair, Smoothing and edge detection by time-varying coupled nonlinear diffusion equations, Computer Vision and Image Understanding 82 (2) (2001) 85–100. doi:10.1006/cviu.2001.0903.

[14] A. Belahmidi, A. Chambolle, Time-delay regularization of anisotropic diffusion and image processing, Mathematical Modelling and Numerical Analysis 39 (2) (2005) 231–251. doi:10.1051/m2an:2005010.

[15] V. B. S. Prasath, D. Vorotnikov, On a system of adaptive coupled pdes for image restoration, Journal of Mathematical Imaging and Vision Online First, available at arXiv:1112.2904. doi:10.1007/s10851-012-0386-3.

[16] D. Strong, Adaptive total variation minimizing image restoration, Ph.D. thesis, UCLA Mathematics Department, USA (August 1997).

[17] Y. Chen, T. Wunderli, Adaptive total variation for image restoration in BV space, Journal of Mathematical Analysis and Applications 272 (3) (2002) 117–137. doi:10.1016/S0022-247X(02)00141-5.

[18] Y. Chen, S. Levine, M. Rao, Variable exponent, linear growth functionals in image restoration, SIAM Journal on Applied Mathematics 66 (4) (2006) 1383–1406. doi:10.1137/050624522.

[19] W. Kusnezow, W. Horn, R. P. Wurtz, Fast image processing with constraints by solving linear PDEs, Electronic Letters on Computer Vision and Image Analysis 6 (2) (2007) 22–35, special Issue: Partial Differential Equations Methods in Graphics and Vision.
URL http://elcvia.uab.es/index.php/elcvia/article/view/146

[20] R. Aboulaich, D. Meskine, A. Souissi, New diffusion models in image processing, Computers & Mathematics with Applications 56 (4) (2008) 874–882. doi:10.1016/j.camwa.2008.01.017.

[21] V. B. S. Prasath, A. Singh, A hybrid convex variational model for image restoration, Applied Mathematics and Computation 215 (10) (2010) 3655–3664. doi:10.1016/j.amc.2009.11.003.

[22] V. B. S. Prasath, A. Singh, Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising, Journal of Applied Mathematics 2010 (2010) 14pp, article ID 763847. doi:10.1155/2010/763847.

[23] P. Guidotti, A new nonlocal nonlinear diffusion of image processing, Journal of Differential Equations 246 (12) (2009) 4731–4742. doi:10.1016/j.jde.2009.03.017.

[24] P. Guidotti, J. Lambers, Two new nonlinear nonlocal diffusions for noise reduction, Journal of Mathematical Imaging Vision 33 (1) (2009) 25–37. doi:10.1007/s10851-008-0108-z.

[25] P. Guidotti, A new well-posed nonlinear nonlocal diffusion, Nonlinear Analysis: Theory, Methods & Applications 72 (12) (2010) 4625–4637. doi:10.1016/j.na.2010.02.040.

[26] P. Guidotti, K. Longo, Two enhanced fourth order diffusion models for image denoising, Journal of Mathematical Imaging Vision 40 (2) (2011) 188–198. doi:10.1007/s10851-010-0256-9.

[27] C. A. Z. Barcelos, M. Boaventura, E. Silva Jr, A well-balanced flow equation for noise removal and edge detection, IEEE Transactions on Image Processing 12 (7) (2003) 751–763. doi:10.1109/TIP.2003.814242.

[28] C. A. Z. Barcelos, M. Boaventura, E. Silva Jr, Edge detection and noise removal by use of a partial differential equation with automatic selection of parameters, Computational & Applied Mathematics 24 (1) (2005) 131–150.

[29] A. I. El-Fallah, G. E. Ford, On mean curvature diffusion in nonlinear image filtering, Pattern Recognition Letters 19 (5–6) (1998) 433–437. doi:10.1016/S0167-8655(98)00030-0.

[30] K. N. Nordstrom, Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection, Image and Vision Computing 8 (4) (1990) 318–327. doi:10.1016/0262-8856(90)80008-H.

[31] B. Smolka, Modified biased anisotropic diffusion processing of noisy color images, in: Ninth IEEE International Conference on Signal Processing, Beijing, China, 2008, pp. 777–780. doi:10.1109/ICOSP.2008.4697245.

[32] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268. doi:10.1016/0167-2789(92)90242-F.

[33] V. Tsurkov, An analytical model of edge protection under noise suppression by anisotropic diffusion, Journal of Computer and Systems Sciences International 39 (3) (2000) 437–440.

[34] S. Kichenassamy, The Perona-Malik method as an edge pruning algorithm, Journal of Mathematical Imaging and Vision 30 (2) (2008) 209–219. doi:10.1007/s10851-007-0029-2.

[35] J. Kaccur, K. Mikula, Slow and fast diffusion effects in image processing, Computing and Visualization in Science 3 (4) (2001) 185–195. doi:10.1007/s007910000047.

[36] V. B. S. Prasath, A well-posed multiscale regularization scheme for digital image denoising, International Journal of Applied Mathematics and Computer Science 21 (4) (2011) 769–777. doi:10.2478/v10006-011-0061-7.

[37] V. B. S. Prasath, A. Singh, An adaptive anisotropic diffusion scheme for image restoration and selective smoothing, International Journal of Image and Graphics 12 (1) (2012) 18pp. doi:10.1142/S0219467812500039.

[38] G. Gilboa, N. Sochen, Y. Y. Zeevi, Variational denoising of partly textured images by spatially varying constraints, IEEE Transactions on Image Processing 15 (8) (2006) 2281–2289. doi:10.1109/TIP.2006.875247.

[39] L. Alvarez, J. Esclarín, Image quantization using reaction-diffusion equations, SIAM Journal on Applied Mathematics 57 (1) (1997) 153–175. doi:10.1137/S0036139994277580.

[40] M. G. Crandall, H. Ishii, P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations, American Mathematical Society Bulletin New Series 27 (1) (1992) 1–67. doi:10.1090/S0273-0979-1992-00266-5.

[41] R. A. Horn, C. R. Johnson, Matrix analysis, Cambridge University Press, Cambridge, UK, 1985.

[42] O. A. Ladyženskaja, V. A. Solonnikov, N. N. Uralʹceva, Linear and quasilinear equations of parabolic type, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., USA, 1967.

[43] L. Alveraz, P. L. Lions, J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion II, SIAM Journal on Numerical Analysis 29 (3) (1992) 845–866. doi:10.1137/0729052.

[44] X. Li, T. Chen, Nonlinear diffusion with multiple edginess thresholds, Pattern Recognition 27 (8) (1994) 1029–1037. doi:10.1016/0031-3203(94)90142-2.

[45] F. Voci, S. Eiho, N. Sugimoto, H. Sekiguchi, Estimating the gradient threshold in the Perona–Malik equation, IEEE Signal Processing Magazine 21 (3) (2004) 39–46. doi:10.1109/MSP.2004.1296541.

[46] J. Weickert, Anisotropic diffusion in image Processing, B.G. Teubner-Verlag, Stuttgart, Germany, 1998.

[47] J. Weickert, Coherence–enhancing diffusion filtering, International Journal of Computer Vision 31 (2–3) (1999) 111–127. doi:10.1023/A:1008009714131.

[48] J. Kaccur, K. Mikula, Slowed anisotropic diffusion, in: First International Conference on Scale-Space Theory in Computer Vision, Vol. 1252, Utrecht, The Netherlands, 1997, pp. 357–360, springer LNCS Eds.: Bart ter Haar Romeny, Luc Florack, Jan Koenderink and Max Viergever. doi:10.1007/3-540-63167-4_71.

[49] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612. doi:10.1109/TIP.2003.819861. URL http://ece.uwaterloo.ca/~z70wang/research/ssim/

[50] C. A. Z. Barcelos, M. Batista, Image restoration using digital inpainting and noise removal, Image and Vision Computing 25 (1) (2007) 61–69. doi:10.1016/j.imavis.2005.12.008.

[51] T. Barbu, V. Barbu, A PDE approach to image restoration problem with observation on a meager domain, Nonlinear Analysis: Real World Applications 13 (3) (2012) 1206–1215. doi:10.1016/j.nonrwa.2011.09.014.