

## Research Article

# Well-Posed Inhomogeneous Nonlinear Diffusion Scheme for Digital Image Denoising

**V. B. Surya Prasath and Arindama Singh**

*Department of Mathematics, Indian Institute of Technology Madras, Chennai 600036, India*

Correspondence should be addressed to Arindama Singh, [asingh@iitm.ac.in](mailto:asingh@iitm.ac.in)

Received 21 August 2009; Revised 13 January 2010; Accepted 11 February 2010

Academic Editor: Malgorzata Peszynska

Copyright © 2010 V. B. S. Prasath and A. Singh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study an inhomogeneous partial differential equation which includes a separate edge detection part to control smoothing in and around possible discontinuities, under the framework of anisotropic diffusion. By incorporating edges found at multiple scales via an adaptive edge detector-based indicator function, the proposed scheme removes noise while respecting salient boundaries. We create a smooth transition region around probable edges found and reduce the diffusion rate near it by a gradient-based diffusion coefficient. In contrast to the previous anisotropic diffusion schemes, we prove the well-posedness of our scheme in the space of bounded variation. The proposed scheme is general in the sense that it can be used with any of the existing diffusion equations. Numerical simulations on noisy images show the advantages of our scheme when compared to other related schemes.

## 1. Introduction

Anisotropic diffusion-based image denoising [1] for gray images was started by Perona and Malik [2] in 1990. It uses a scale parameter-based adaptive image filtering using nonlinear diffusion. Let  $u_0$  be a noisy version of the true image  $u$  with noise field  $n$  of known variance  $\sigma^2$ :

$$u_0 = u + n. \quad (1.1)$$

Our aim is to recover the noise-free image  $u$  with edge preservation. To avoid the oversmoothing nature of the linear diffusion, Perona and Malik [2] introduced an edge indicator function  $g$  for reducing the diffusion near edges via the following anisotropic diffusion scheme (ADS):

$$\frac{\partial u(x, t)}{\partial t} = \operatorname{div}(g(|\nabla u(x, t)|)\nabla u(x, t)), \quad \text{for } x \in \Omega \quad (1.2)$$

with the initial condition  $u(x, 0) = u_0(x)$ . The diffusion function  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  is chosen to be decreasing, such that  $g(0) = 1$ ,  $\lim_{s \rightarrow \infty} g(s) = 0$ . If  $g(s) \equiv 1$  in (1.2), then we recover the linear diffusion PDE. The diffusion functions  $g(\cdot)$  proposed in [2] are

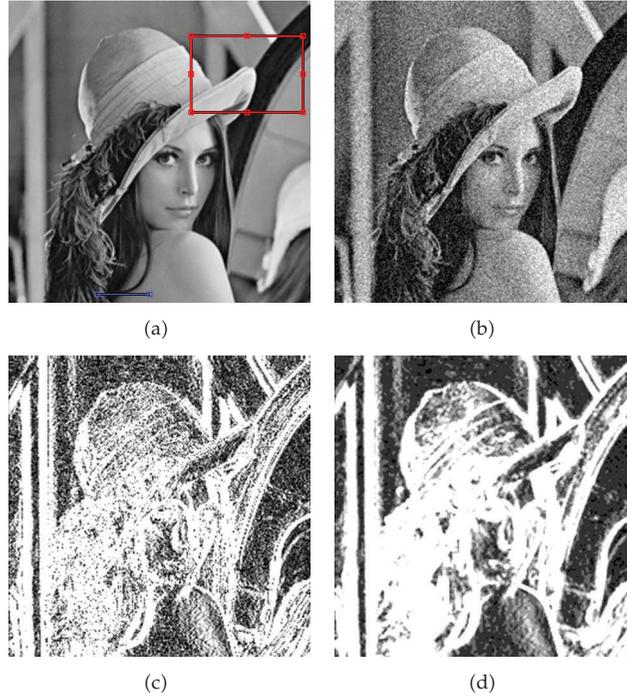
$$g_1(s) = \exp\left(-\left(\frac{s}{K}\right)^2\right), \quad g_2(s) = \frac{1}{1 + (s/K)^2}, \quad (1.3)$$

where  $K > 0$  is the so-called contrast parameter. Better numerical results are obtained in [2] using this class of anisotropic PDEs (1.2)-(1.3) instead of the linear diffusion. When  $|\nabla u| > K$ , the PDE (1.2) turns into inverse diffusion, which is known to be ill-posed [3]. It is easily seen that the role of  $g$  as an edge indicator function is important in reducing the noise and enhancing true edges. A more robust indicator function based on smoothed gradient,  $g(|\nabla G_\rho \star u|)$ , where the Gaussian kernel of width  $\rho$  is given by  $G_\rho(x) = (2\pi\rho^2)^{-1} \exp(-|x|^2/2\rho^2)$ , can be used for edge identification. This presmoothing also alleviates the instability [4] associated with the ADS (1.2). However, edge localization is lost and noisy oscillations can remain in this method. Moreover, when the noise level is high, ADS can reduce the contrast of edges and reliance on the instable  $|\nabla u|$  alone can be a drawback.

There are efforts to remedy the ill-posedness and to use better edge indicators in the diffusion process. Strong [5] used an adaptive parameter in total variation-based PDE using the relationship between energy minimization and PDEs [3]. Better results are obtained when compared to the classical total variation (TV) scheme of [6, 7], but the use of gradients alone and the noisy edge map provided by  $u_0$  can lead to noise amplification along edges. Recently, Ceccarelli et al. [8] devised well-posed ADS schemes by approximating the TV function. Kuznezov et al. [9] used a similar weight function of Strong [5] combined with linear diffusion for fast computations. Yu et al. [10] considered ADS in terms of kernel-based smoothing and proposed to use a diffusion function based on modified gradients. Barbu et al. [11] considered the variational-PDE problem in Sobolev space setting with different growth functions. Douiri et al. [12] use an edge-based weight in the regularization functional for diffuse optical tomography to control smoothing across edges. The exact edge information is used to compute the weight which in real images is not possible, since we do not know, a priori, the exact edge map of the true image. Indeed, this edge estimation problem is closely related to image denoising, and most of the edge detection schemes [13, 14] are based on this observation.

Apart from staircasing artifacts in flat regions and noise amplification along edges, capturing multiscale edges are difficult in previous schemes. To circumvent the drawbacks a more stable and robust edge indicator function using the multiscale edge detectors as discussed in [13–15] can be used. Anisotropic diffusion scheme is usually employed as a preprocessing step, for noise removal and edge detection. Here, we reverse the trend and integrate the edge detection into the anisotropic scheme such as (1.2). The method can, in general, be used in any anisotropic PDE, and it provides a better coupling of edge detection and restoration of images under noise. We prove the well-posedness of the proposed scheme in the space of bounded variation  $BV(\Omega)$  [16]. The analysis is based on monotone operators and approximation schemes. The discrete version is implemented and it satisfies the properties required for an edge preserving scale space. Preliminary numerical results were reported in [17].

The rest of the paper is organized as follows. In Section 2, we introduce the edge adaptive PDE scheme and in Section 3 its well-posedness is proved in the space of bounded variation functions. We show the numerical results on real and noisy images in Section 4. Finally, Section 5 concludes the paper.



**Figure 1:** Image gradients based edge indicators. (a) Original image. (b) Noisy image. (c)  $g_1(|\nabla u|)$ . (d)  $g_1(|G_\sigma * u|)$ .

## 2. Proposed Scheme

### 2.1. Motivation

The diffusion function  $g$  in ADS (1.2) uses the information provided by magnitude of the gradient to reduce diffusion near edges. Figure 1 shows the effect of this on *Lena* test image. Original image (Figure 1(a)) is corrupted with additive Gaussian noise of variance 20 and is shown in Figure 1(b). ADS (1.2) uses only the diffusion coefficient  $g$  as in (1.3). ADS restores the image in a piecewise constant manner but reduces overall contrast and creates staircasing artifacts near edges; see Figure 4(a). The magnitude of the gradient image (Figure 1(c), where whiter pixels represent maximal value of  $|\nabla u|$ ) shows that it is not reliable under noise, and loss of spatial coherence such as edge connectivity and continuity is high. Smoothed gradient ( $|\nabla G_\rho * u|$ ) method of [4] reduces this loss but still leaves some spurious pixels due to noise (Figure 1(d)) and true locations of the edges are not preserved. This makes the result lacking clear boundaries; see Figure 4(b). Also this use of isotropic diffusion is against the very principle of ADS (1.2), which is anisotropic in nature.

### 2.2. Edge Adaptive ADS

Instead of using only the edge indicator function  $g$  alone to drive the diffusion process, we introduce a spatially adaptive term  $\alpha$  in the anisotropic scheme ( $\alpha$ -ADS):

$$\frac{\partial u(x, t)}{\partial t} = \operatorname{div}(\alpha(x, t)g(|\nabla u(x, t)|)\nabla u(x, t)) - \lambda(u - u_0), \quad (2.1)$$

where  $\alpha$  provides a pixelwise edge characterization. As we will see in Section 3 and in the numerical experiments (Section 4) this avoids the stability problems of the gradient magnitude-based schemes and overcomes the localization error of smoothed gradient method. The requirements for  $\alpha$  are as follows:

- (1) at edge pixels to reduce the diffusion:  $\alpha(x, t) \sim 0$  on the edges,
- (2) in flat regions to allow smoothing:  $\alpha(x, t) \sim 1$  beyond the edges,
- (3) near edge pixels  $\alpha(x, t)$  should vary smoothly between 0 and 1 so as to avoid spurious oscillations.

We choose a term of the form

$$\alpha(x, t) = 1 - E(x, t) \quad (2.2)$$

with  $E$  as a robust and smooth edge indicator function.

### 2.3. Choice of Edge Indicator Function

Various edge detectors [13–15] exist in digital image processing literature. These detectors are based on a principle of gradient maxima. Among the available edge detectors [14] such as Sobel's, Prewitt's, Laplacian of Gaussian (LOG), and Marr-Hildreth's Zero-crossing that use maxima points of gradient to decide about the edge pixels, Canny's edge detector [13, 15] has been proved to be the best in terms of spatial coherency and edge continuity. The general method of Canny involves the following steps:

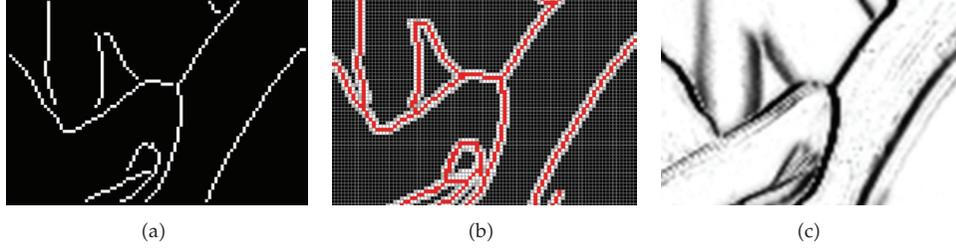
- (1) convolution of the given image  $u$  by a Gaussian  $G_\sigma$ ,
- (2) estimating the second derivative  $\nabla^2(u)$  using finite differences,
- (3) again using convolution for  $\nabla^2(u)$  with a small Gaussian kernel,
- (4) thresholding the gradient of Step 1,
- (5) zero-crossings of Step 3 displayed if threshold of Step 4 is achieved.

Canny edge map  $C(x, t)$  is a binary output with edge pixels marked as 1 and nonedge pixels with 0. Hence we use

$$E(x, t) = G_\rho \star C(x, t), \quad (2.3)$$

where  $G_\rho$  smoothens the transition from detected edges to homogenous regions. Figure 2 shows a closeup of the edge map from the hat region of *Lena* image given by a rectangle in Figure 1(b). Clearly the Canny edge detector-based smoothing parameter  $E(x, t)$  gives a good localization of edges (Figure 2(b)) and is devoid of noisy pixels. Compare this with the  $|\nabla u|$  or the smoothed version  $|\nabla G_\sigma \star u|$ -based edge indicators (Figures 1(c) and 1(d)) where spatial coherency is missing and edge continuity is lost. Also, in the diffusion process based on these measures, the noisy pixels found in homogenous regions are propagated resulting in staircase artifacts.

One can use any of the mentioned edge detectors into the spatially adaptive term  $\alpha(x, t)$  in (2.2). We choose Canny detector because of its superiority and its computational efficiency. This trade-off between optimal performance and computation time can be decided according to the problem at hand.



**Figure 2:** Canny edge detector-based indicator function. (a) Canny binary output  $C(x)$ . (b) Smooth edge indicator function  $E(x)$  with  $\rho = 2$  creates a band around  $C(x)$  (c) Diffusion weight  $\alpha g(|\nabla u|)$ .

### 3. Existence and Uniqueness

Let  $\Omega \subset \mathbb{R}^2$  be a bounded domain. The proposed spatially adaptive version  $\alpha$ -ADS (2.1) can be considered as a gradient descent form of a regularization functional [3]. The proposed scheme (2.1) is equivalent to a descent method for solving the following energy minimization problem:

$$\min_{u \in \text{BV}(\Omega)} \left\{ \mathcal{E}(u) = \frac{\lambda}{2} \int_{\Omega} |u - u_0|^2 dx + \int_{\Omega} \alpha \phi(|\nabla u|) dx \right\}, \quad (3.1)$$

with  $\phi'(s) = g(s)s$ . For example, for the diffusion functions in (1.3) we get the regularization functions as

$$\phi_1(s) = 1 - \exp\left(-\left(\frac{s}{K}\right)^2\right), \quad \phi_2(s) = \log\left(1 + \left(\frac{s}{K}\right)^2\right). \quad (3.2)$$

Since these two functions are nonconvex, the minimization problem (3.1) and ADS (1.2) do not have a unique solution. In these nonconvex regularization function cases the corresponding energy minimization problem (3.1) does not have a solution at all, except when  $u = u_0 = \text{constant}$ ; see [3] for an analysis between the PDE (1.2) and minimization problem (3.1). To obtain a well-posed scheme we restrict ourselves to the class of linear growth functions for  $\phi : \mathbb{R} \rightarrow [0, \infty)$  satisfying the following.

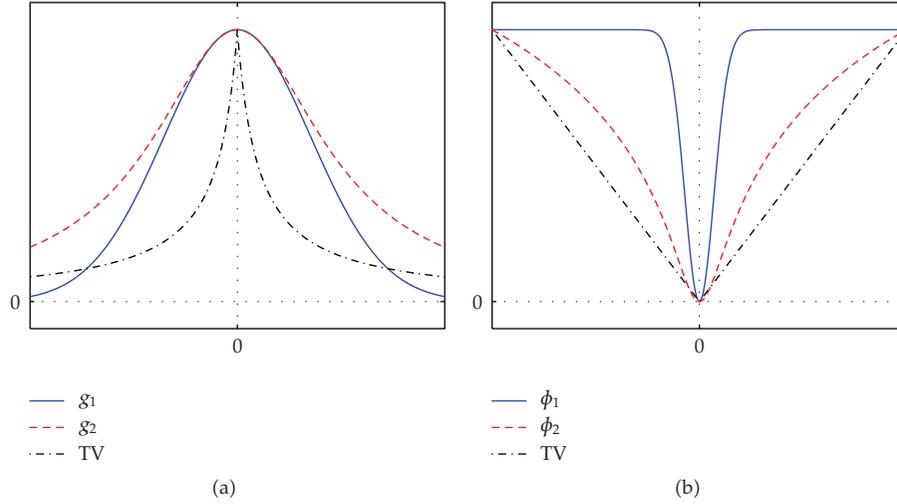
(H1)  $\phi$  is a nondecreasing, convex, and even function with  $\phi(0) = 0$ .

(H2) There exist  $a > 0$  and  $b \geq 0$  such that  $as - b \leq \phi(s) \leq as + b$ , for all  $s$ .

An important example in this class is the minimal surface function  $\phi(s) = \sqrt{1 + s^2}$ , which is related to the total variation (TV) regularization [6, 7]; see Figure 3. We assume the preliminary results about the space of functions of bounded variation  $\text{BV}(\Omega)$ ; see [16]. Throughout this paper, we use the notation  $\phi(x, \nabla u) = \alpha(x)\phi(|\nabla u|)$ . We recall the definition of total variation of a function  $u \in L^1(\Omega)$ .

*Definition 3.1.* Let  $u \in L^1(\Omega)$ ; its total variation is defined as

$$\int_{\Omega} |\nabla u| := \sup_{|w| \leq 1} \left\{ \int_{\Omega} u \operatorname{div} w dx : w \in C_0^1(\Omega) \right\}. \quad (3.3)$$



**Figure 3:** (a) Diffusion functions  $g_1, g_2$  given in (1.3) and TV function. (b) Corresponding regularization functions.

The total variation is in fact a Radon measure; see [1, page 50] for more details. Making use of the properties of the convex functions of measures from [18] we will prove that the  $\alpha$ -ADS (2.1) has a unique weak solution in the following sense.

*Definition 3.2.*  $u \in L^2(\text{BV}(\Omega); [0, T])$  is a weak solution of the PDE (2.1) if for every  $v \in L^2(\text{BV}(\Omega); [0, T])$ , and  $\tau \in [0, T]$

$$\begin{aligned} & \int_0^\tau \int_\Omega \phi(x, \nabla u) dx dt - \lambda \int_0^\tau \int_\Omega (u - u_0)(v - u) dx dt \\ & \leq \int_0^\tau \int_\Omega \frac{\partial u}{\partial t} (v - u) dx dt + \int_0^\tau \int_\Omega \phi(x, \nabla v) dx dt, \quad \text{a.e. in } \tau. \end{aligned} \quad (3.4)$$

Moreover, it satisfies  $\partial u / \partial t \in L^2(\Omega \times [0, T])$  and  $u(x, 0) = u_0(x)$ .

The function space  $L^2(\text{BV}(\Omega); [0, T])$  is the set of all functions  $w : \Omega \times [0, T] \rightarrow \mathbb{R}$  such that,  $w(\cdot, t) \in \text{BV}(\Omega)$  for each  $t$  and  $w(x, \cdot) \in L^2([0, T])$  for all  $x$ . We consider the following regularized linear growth function (see [7]) in the approximation problem of (2.1):

$$\phi_\epsilon(x, s) := \begin{cases} \frac{1}{2\epsilon} \frac{\partial \phi(x, s)}{\partial s} \Big|_{s=\epsilon} s^2 - \frac{\epsilon}{2} \frac{\partial \phi(x, s)}{\partial s} \Big|_{s=\epsilon} + \phi(x, \epsilon), & \text{if } s < \epsilon, \\ \phi(x, s), & \text{if } \epsilon \leq s \leq \frac{1}{\epsilon}, \\ \frac{\epsilon}{2} \frac{\partial \phi(x, s)}{\partial s} \Big|_{s=1/\epsilon} s^2 - \frac{1}{2\epsilon} \frac{\partial \phi(x, s)}{\partial s} \Big|_{s=\epsilon} + \phi(x, 1/\epsilon), & \text{if } s > \frac{1}{\epsilon}. \end{cases} \quad (3.5)$$

**Lemma 3.3.** *Let the regularization function  $\phi$  satisfy the assumptions H1 – H2. If  $\{u_i\}$  converges to  $u$  in  $L^1(\Omega)$ , then*

$$\int_{\Omega} \phi(x, \nabla u) dx \leq \liminf_{i \rightarrow \infty} \int_{\Omega} \phi(x, \nabla u_i) dx. \quad (3.6)$$

*Proof.* Let  $w \in C_0^1(\Omega)$  be such that  $|w(x)| \leq \alpha(x)$  for all  $x \in \Omega$ . Then

$$\int_{\Omega} u \operatorname{div} w dx = \lim_{i \rightarrow \infty} \int_{\Omega} u_i \operatorname{div} w dx \leq \liminf_{i \rightarrow \infty} \int_{\Omega} \alpha |\nabla u_i| dx, \quad (3.7)$$

where the inequality follows from Definition 3.1 of total variation of  $u$ . Taking supremum over  $w$  gives the result.  $\square$

Note that  $\alpha$  defined in (2.2) is a smooth function, that is,  $\alpha(x) \in C^\infty(\Omega; (0, \infty))$  due to the Gaussian presmoothing; see (2.3). The above lemma can also be proved using a general assumption about the adaptive parameter such as  $\alpha \in C^0(\Omega; (0, \infty))$ ; see also [19, Theorem 1]. We next need the following result from nonlinear semigroup theory; see Theorem 3.1 in [20].

**Theorem 3.4.** *Let  $H$  be a Hilbert space, and let  $P(H)$  be its power set. Let  $\mathcal{A} : H \rightarrow P(H)$  be a maximal monotone operator and  $u_0 \in D(\mathcal{A})$ , where the domain is  $D(\mathcal{A}) = \{x \in H : \mathcal{A}x \neq \emptyset\}$ . Then there exists a unique function  $u(t) : [0, \infty) \rightarrow H$  such that*

$$0 \in \frac{\partial u}{\partial t} + \mathcal{A}(u), \quad u(0) = u_0. \quad (3.8)$$

**Lemma 3.5.** *Consider the approximation PDE*

$$\begin{aligned} -\epsilon \Delta u + \frac{\partial u}{\partial t} - \operatorname{div}(g_\epsilon(x, |\nabla u|) \nabla u) + \lambda(u - u_0^\delta) &= 0 \quad \text{on } \Omega \times [0, T], \\ u &= u_0^\delta \quad \text{on } \Omega \times \{0\} \end{aligned} \quad (3.9)$$

with

$$\begin{aligned} g_\epsilon(x, \xi) &= \frac{1}{\xi} \frac{\partial \phi_\epsilon(x, \xi)}{\partial \xi}, \quad u_0^\delta \in C^\infty(\overline{\Omega}), \\ u_0^\delta &\rightarrow u_0 \text{ in } L^2(\Omega) \quad \text{with } \|u_0^\delta\|_{L^\infty(\Omega)} \leq C \|u_0\|_{L^\infty(\Omega)}. \end{aligned} \quad (3.10)$$

There exists a unique solution  $u_\epsilon^\delta \in L^\infty(W^{1,2}(\Omega); [0, T])$  for the approximation PDE. The solution  $u_\epsilon^\delta$  is bounded in  $L^\infty(\Omega)$  by the initial value  $u_0$ , that is,

$$\|u_\epsilon^\delta\|_{L^\infty(\Omega \times [0, T])} \leq C \|u_0\|_{L^\infty(\Omega)}. \quad (3.11)$$

*Proof.* Since  $u_0 \in \text{BV}(\Omega) \cap L^\infty(\Omega)$ , existence of the sequence  $u_0^\delta$  is guaranteed. Let  $\mathcal{A}_\epsilon(u) = \lambda(u - u_0^\delta) - \text{div}(g_\epsilon(x, |\nabla u|)\nabla u)$ . As  $\mathcal{A}_\epsilon$  is the derivative of a convex lower semicontinuous functional, it is a maximal monotone operator. The existence of  $u_\epsilon^\delta$  follows from Theorem 3.4.

To get the  $L^\infty$  bound for the solution we proceed as follows. Let  $B = \|u_0^\delta\|_{L^\infty(\Omega)}$  and  $f^+ = \max(f, 0)$ . Multiply the approximation PDE by the term  $e^{-st}(e^{-st}u_\epsilon^\delta - B)^+$ . Integrating the PDE, we obtain

$$\begin{aligned} & \epsilon \int_{\Omega} \nabla u_\epsilon^\delta e^{-st} \cdot e^{-st} \nabla u_\epsilon^\delta dx + \int_{\Omega} \frac{\partial u_\epsilon^\delta}{\partial t} e^{-st} (e^{-st} u_\epsilon^\delta - B)^+ dx \\ & + \int_{\Omega} [g_\epsilon(x, |\nabla u_\epsilon^\delta|) \nabla u_\epsilon^\delta] e^{-st} \cdot e^{-st} \nabla u_\epsilon^\delta dx \\ & + \lambda \int_{\Omega} (u_\epsilon^\delta - u_0^\delta) e^{-st} (e^{-st} u_\epsilon^\delta - B)^+ dx = 0. \end{aligned} \quad (3.12)$$

This implies that

$$\int_{\Omega} \frac{\partial u_\epsilon^\delta}{\partial t} e^{-st} (e^{-st} u_\epsilon^\delta - B)^+ dx \leq 0 \quad (3.13)$$

since the remaining integrals are all nonnegative. Let

$$\mathcal{J}(t) = \int_{\Omega} \left| (e^{-st} u_\epsilon^\delta - B)^+ \right|^2 dx. \quad (3.14)$$

Then  $\mathcal{J}(t)$  is decreasing, nonnegative, and  $\mathcal{J}(0) = 0$ . Thus  $\mathcal{J}(t) = 0$  for all  $t$ . Subsequently,

$$u_\epsilon^\delta(t) \leq B e^{st} \quad \text{a.e. on } \Omega \text{ for } s, t > 0. \quad (3.15)$$

Letting  $s \rightarrow 0$ , we get  $u_\epsilon^\delta(t) \leq \|u_0^\delta\|_{L^\infty(\Omega)}$ . Similarly multiplying the term  $e^{-st}(-B - e^{-st}u_\epsilon^\delta)^+$  with the approximation PDE and integrating, we obtain  $u_\epsilon^\delta(t) \geq -\|u_0^\delta\|_{L^\infty(\Omega)}$ .  $\square$

**Lemma 3.6.** *The weak solution of the approximation PDE  $u_\epsilon^\delta$  satisfied the following inequality, for  $\tau \in [0, T]$ :*

$$\begin{aligned} & \int_0^\tau \int_{\Omega} \left( \frac{\partial u_\epsilon^\delta}{\partial t} \right)^2 dx dt + \frac{\epsilon}{2} \int_{\Omega} |\nabla u_\epsilon^\delta(t)|^2 dx + \int_{\Omega} \phi(x, \nabla u_\epsilon^\delta(t)) dx \\ & \leq \frac{\epsilon}{2} \int_{\Omega} |\nabla u_0^\delta|^2 dx + \int_{\Omega} \phi(x, \nabla u_0^\delta) dx + \epsilon, \quad \text{a.e. in } \tau. \end{aligned} \quad (3.16)$$

Moreover, the weak solutions  $u_\epsilon^\delta$  and  $\partial u_\epsilon^\delta / \partial t$  are uniformly bounded in  $\epsilon$  with respect to the  $L^\infty(\Omega \times [0, T])$  and  $L^2(\Omega \times [0, T])$  norms, respectively.

*Proof.* The proof of the inequality follows from the identity

$$\begin{aligned} \int_0^T \int_{\Omega} \left( \frac{\partial u_{\epsilon}^{\delta}}{\partial t} \right)^2 dx dt + \frac{\epsilon}{2} \int_{\Omega} |\nabla u_{\epsilon}^{\delta}(t)|^2 dx + \int_{\Omega} \phi_{\epsilon}(x, \nabla u_{\epsilon}^{\delta}(t)) dx + \frac{\lambda}{2} \int_{\Omega} (u_{\epsilon}^{\delta} - u_0^{\delta})^2 dx \\ = \frac{\epsilon}{2} \int_{\Omega} |\nabla u_0^{\delta}|^2 dx + \int_{\Omega} \phi_{\epsilon}(x, \nabla u_0^{\delta}) dx \end{aligned} \quad (3.17)$$

and  $\phi(x, s) \leq \phi_{\epsilon}(x, s) \leq \phi(x, s) + \epsilon$  for any  $s$ .

From Lemma 3.5, it follows that  $u_{\epsilon}^{\delta}$  is uniformly bounded and

$$\int_0^T \int_{\Omega} \left( \frac{\partial u_{\epsilon}^{\delta}}{\partial t} \right)^2 dx dt + \int_{\Omega} \phi(x, \nabla u_{\epsilon}^{\delta}) dx \leq \frac{\epsilon}{2} \int_{\Omega} |\nabla u_0^{\delta}|^2 dx + \int_{\Omega} \phi(x, \nabla u_0^{\delta}) dx + \epsilon. \quad (3.18)$$

We have

$$\int_0^T \int_{\Omega} \left( \left| \frac{\partial u_{\epsilon}^{\delta}}{\partial t} \right| + \phi(\nabla u_{\epsilon}^{\delta}) \right) dx dt \leq C \|u_0^{\delta}\|_{L^{\infty}(\Omega)}. \quad (3.19)$$

□

Let  $\{\epsilon_n\}$  be a sequence converging to 0. Consider the corresponding weak solutions of the approximation PDE  $\{u_{\epsilon_n}^{\delta}\}$ . In the following lemma, we prove the convergence of a subsequence of weak solutions  $\{u_{\epsilon_{n_k}}^{\delta}\}$  as  $\epsilon_{n_k} \rightarrow 0$ .

**Lemma 3.7.** *There exists a subsequence  $\{u_{\epsilon_{n_k}}^{\delta}\}$  such that as  $\epsilon_{n_k} \rightarrow 0$*

$$\begin{aligned} u_{\epsilon_{n_k}}^{\delta} &\longrightarrow u^{\delta} \quad \text{strongly in } L^1(\Omega \times [0, T]), \quad \text{a.e. in } \Omega \times [0, T], \\ \frac{\partial u_{\epsilon_{n_k}}^{\delta}}{\partial t} &\rightharpoonup \frac{\partial u^{\delta}}{\partial t} \quad \text{weakly in } L^2(\Omega \times [0, T]), \end{aligned} \quad (3.20)$$

where  $u^{\delta}$  is a weak solution of the PDE (2.1) with  $u^{\delta}(0) = u_0^{\delta}$ . Moreover  $\|u^{\delta}\|_{L^{\infty}(\Omega \times [0, T])} \leq C \|u_0\|_{L^{\infty}(\Omega)}$ .

*Proof.* By Lemma 3.6,  $u_{\epsilon}^{\delta}$  and  $\partial u_{\epsilon}^{\delta} / \partial t$  are uniformly bounded. When  $\epsilon_{n_k} \rightarrow 0$ , for fixed  $\delta > 0$ , we have a convergent subsequence  $u_{\epsilon_{n_k}}^{\delta} \rightarrow u^{\delta}$  in  $L^1(\Omega \times [0, T])$  for  $u^{\delta} \in L^{\infty}(\text{BV}(\Omega) \cap L^{\infty}(\Omega); [0, T])$ , and  $(\partial u_{\epsilon_{n_k}}^{\delta} / \partial t) \rightharpoonup (\partial u^{\delta} / \partial t)$  in  $L^2(\Omega \times [0, T])$ .

Since  $u_{\epsilon_{n_k}}^{\delta}$  is the weak solution of the approximation PDE, it satisfies the corresponding weak solution formulation in Definition (3.4). That is, for every  $v \in L^2(\text{BV}(\Omega); [0, T])$  and  $\tau \in [0, T]$ ,

$$\begin{aligned} \int_0^{\tau} \int_{\Omega} \phi_{\epsilon_{n_k}}(x, \nabla u_{\epsilon_{n_k}}^{\delta}) dx dt - \lambda \int_0^{\tau} \int_{\Omega} (u_{\epsilon_{n_k}}^{\delta} - u_0^{\delta}) (v - u_{\epsilon_{n_k}}^{\delta}) dx dt \\ \leq \frac{\epsilon_{n_k}}{2} \int_0^{\tau} \int_{\Omega} |\nabla v|^2 dx dt + \int_0^{\tau} \int_{\Omega} \phi_{\epsilon_{n_k}}(x, \nabla v) dx dt + \int_0^{\tau} \int_{\Omega} \frac{\partial u_{\epsilon_{n_k}}^{\delta}}{\partial t} (v - u_{\epsilon_{n_k}}^{\delta}) dx dt \quad \text{a.e. in } \tau. \end{aligned} \quad (3.21)$$

When  $\epsilon_{n_k} \rightarrow 0$ , we obtain, using Lemma 3.3,

$$\begin{aligned} & \int_0^\tau \int_\Omega \phi(x, \nabla u^\delta) dx dt - \lambda \int_0^\tau \int_\Omega (u^\delta - u_0^\delta)(v - u^\delta) dx dt \\ & \leq \int_0^\tau \int_\Omega \phi(x, \nabla v) dx dt + \int_0^\tau \int_\Omega \frac{\partial u^\delta}{\partial t} (v - u^\delta) dx dt. \end{aligned} \quad (3.22)$$

Thus,  $u^\delta$  is a weak solution of the PDE (2.1) with  $u^\delta(x, 0) = u_0^\delta(x)$ .  $\square$

Now, we state and prove our main result.

**Theorem 3.8.** *If  $u_0 \in BV(\Omega) \cap L^\infty(\Omega)$ , then there exists a unique weak solution  $u \in L^\infty(BV(\Omega) \cap L^\infty(\Omega); [0, T])$  of (2.1).*

*Proof.* From Lemma 3.7 we get  $u^\delta$  as a weak solution of the PDE (2.1) with  $u^\delta(x, 0) = u_0^\delta(x)$ . Using Lemma 3.7 for the sequence  $\{u^\delta\}$  as  $\delta \rightarrow 0$ , we get

$$\begin{aligned} u^\delta & \longrightarrow u \quad \text{strongly in } L^1(\Omega \times [0, T]), \quad \text{a.e. in } \Omega \times [0, T], \\ \frac{\partial u^\delta}{\partial t} & \rightharpoonup \frac{\partial u}{\partial t} \quad \text{weakly in } L^2(\Omega \times [0, T]). \end{aligned} \quad (3.23)$$

Finally we let  $\delta \rightarrow 0$  in the following inequality:

$$\begin{aligned} & \int_0^\tau \int_\Omega \phi(x, \nabla u^\delta) dx dt - \lambda \int_0^\tau \int_\Omega (u^\delta - u_0^\delta)(v - u^\delta) dx dt \\ & \leq \int_0^\tau \int_\Omega \phi(x, \nabla v) dx dt + \int_0^\tau \int_\Omega \frac{\partial u^\delta}{\partial t} (v - u^\delta) dx dt \end{aligned} \quad (3.24)$$

to obtain

$$\begin{aligned} & \int_0^\tau \int_\Omega \phi(x, \nabla u) dx dt - \lambda \int_0^\tau \int_\Omega (u - u_0)(v - u) dx dt \\ & \leq \int_0^\tau \int_\Omega \phi(x, \nabla v) dx dt + \int_0^\tau \int_\Omega \frac{\partial u}{\partial t} (v - u) dx dt. \end{aligned} \quad (3.25)$$

We see that  $u \in L^\infty(BV(\Omega) \cap L^\infty(\Omega); [0, T])$  is a weak solution of (2.1). Uniqueness of the solution follows from the weak solution inequality (3.4) above.  $\square$

As a consequence of Theorem 3.8, we also obtain a bound for the solution  $\|u\|_{L^\infty} \leq C\|u_0\|_{L^\infty}$ , that is, the maximum principle holds for the solution of (2.1); this guarantees that no new structures are created and  $\alpha$ -ADS (2.1) satisfies the scale space axiom. Note that though  $\alpha \in C^\infty(\Omega)$  (since  $G_\rho \in C^\infty(\Omega)$ ) in (2.3), if we use the original ADS functions (1.3) we cannot obtain the well-posedness of the PDE (2.1), since the corresponding  $\phi$  functions are not convex.

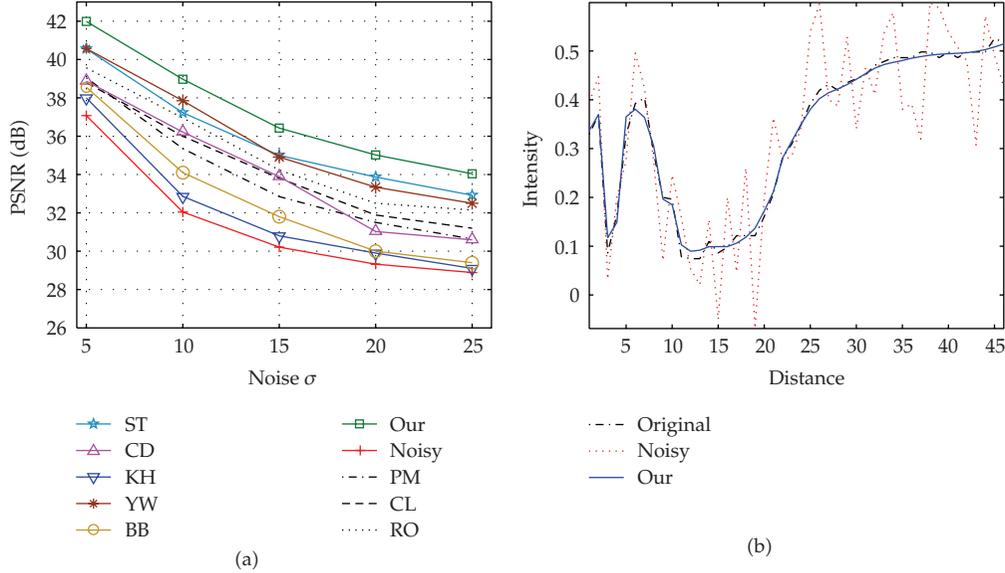


**Figure 4:** Results of (a) Perona and Malik scheme [2], (b) Catté et al. scheme [4], (c) Rudin et al. scheme [6], (d) Strong scheme [5], (e) Ceccarelli et al. scheme [8], (f) Kusnezow et al. scheme [9], (g) Yu et al. scheme [10], (h) Barbu et al. scheme [11], (i) Our  $\alpha$ -ADS scheme.

#### 4. Numerical Experiments

We use finite differences to discretize the proposed edge adaptive PDE (2.1) for images normalized to be in the range  $[0, 1]$ . We take  $h$  as the grid size, and  $U_{ij}^t$  as the intensity value  $u(i, j)$  at iteration  $t$ . Instead of the classical explicit scheme, which severely restricts the step size, we make use of the unconditionally stable semi-implicit scheme. In 1D with matrix-vector notation it reads as

$$U^{t+1} = [1 - \tau A(U^t)]^{-1} U^t, \quad (4.1)$$



**Figure 5:** (a) Noise level  $\sigma$  versus PSNR (dB) values. (b) Profile line taken across *Lena* at pixel position 250.

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

$$a_{ij}(U^t) := \begin{cases} \frac{\gamma_i^t + \gamma_j^t}{2h^2}, & j \in \mathcal{N}_i, \\ -\sum_{k \in \mathcal{N}_i} \frac{\gamma_i^t + \gamma_k^t}{2h^2}, & j = i, \\ 0, & \text{otherwise} \end{cases} \quad (4.2)$$

with  $\gamma_i = \alpha_i g_i$  and  $\mathcal{N}_i$  is the discrete neighborhood of the pixel  $i$ . For n-D images the semi-implicit scheme is written as

$$U^{t+1} = \left[ 1 - \tau \sum_{l=1}^n A_l(U^t) \right]^{-1} U^t. \quad (4.3)$$

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

We use the Canny edge detector with its default settings for (2.3) in MATLAB7.4, and the computations are done in a desktop computer with Pentium IV, 2.45 GHz processor. It took nearly a minute for 100 iterations of our scheme, since the computation of Canny edge detector is done at each iteration. Further reduction in execution time can be achieved if we are to use other faster edge detectors. The numerical example, given in Figure 4, shows the results obtained for the noisy *Lena* image given in Figure 1(b). We show results of applying

our scheme (2.1) with regularized TV function  $g(s) = (\epsilon^2 + s^2)^{-1/2}$ ,  $\epsilon = 10^{-6}$  in Figure 4(i) and also compare with the following schemes:

- (I) *classical schemes*: ADS (1.2) of Perona and Malik [2] (Figure 4(a)), Catté et al.'s smoothed gradient scheme [4] (Figure 4(b)), and Rudin et al.'s total variation scheme [6] (Figure 4(c)),
- (II) *adaptive schemes*: Strong's adaptive total variation scheme [5] (Figure 4(d)), Ceccarelli et al.'s approximate TV scheme [8] (Figure 4(e)), Kusnezow et al.'s adaptive linear diffusion scheme [9] (Figure 4(f)), Yu et al.'s kernel-based ADS [10] (Figure 4(g)), and Barbu et al.'s variational PDE scheme [11] (Figure 4(h)).

We use the peak signal-to-noise ratio (PSNR) for comparing our scheme with other schemes. The parameters are tuned to get the best possible PSNR value for each of the scheme compared. Figure 5(a) shows the effect of the same against different noise levels ( $5 \leq \sigma \leq 25$ ) for our scheme. For an  $m \times n$  image  $u$ , the PSNR value of an estimated image  $\hat{u}$  is given by

$$\text{PSNR} = 10 \log_{10} \left( \frac{m \times n}{\sqrt{\sum_{x \in \Omega} (u(x) - \hat{u}(x))^2}} \right) (dB). \quad (4.4)$$

We see that our scheme produces a stable PSNR value as the noise level increases and attains highest values among other related schemes as well. To show the strong smoothing along flat regions and edge preservation, Figure 5(b) shows a signal line taken across *Lena* image from Figure 4(i), whose position is indicated in Figure 1(a) by a line. As can be seen our scheme can remove small scale texture details as the edge detector in (2.2) is not sufficient to capture them. Incorporation of textural measures into  $\alpha$  points at further improvements in this direction.

## 5. Conclusions

A class of well-posed inhomogeneous diffusion schemes for image denoising is studied in this paper. By integrating multiscale edge detectors into the divergence term of anisotropic diffusion PDE we obtain edge preserving restoration of noisy images. Unlike the classical anisotropic diffusion schemes, which use only gradients to find edge pixels, we have proposed to include an adaptive parameter computed from Canny edge detector. Using an approximation scheme and the theory of monotone operators, well-posedness of the proposed inhomogeneous PDE is proved in the space of functions of bounded variation. Numerical examples illustrate that the scheme performs well under noise. Comparison with related schemes is undertaken to show that the expected improvement can really be achieved. Similar analysis for higher-order diffusion schemes and comparing various edge detectors for performance evaluation provide future directions.

## Acknowledgment

The authors thank the anonymous reviewers for their comments which improved the content and presentation of the paper and the handling editor Professor Malgorzata Peszynska for her patience during the review period.

## References

- [1] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2nd edition, 2006.
- [2] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [3] Y.-L. You, W. Xu, A. Tannenbaum, and M. Kaveh, "Behavioral analysis of anisotropic diffusion in image processing," *IEEE Transactions on Image Processing*, vol. 5, no. 11, pp. 1539–1553, 1996.
- [4] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.
- [5] D. Strong, *Adaptive total variation minimizing image restoration*, Ph.D. thesis, UCLA Mathematics Department, Los Angeles, Calif, USA, 1997, <ftp://ftp.math.ucla.edu/pub/camreport/cam97-38.ps.gz>.
- [6] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [7] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.
- [8] M. Ceccarelli, V. De Simone, and A. Murli, "Well-posed anisotropic diffusion for image denoising," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 149, no. 4, pp. 244–252, 2002.
- [9] W. Kusnezow, W. Horn, and R. P. Würtz, "Fast image processing with constraints by solving linear PDEs," *Electronic Letters in Computer Vision and Image Analysis*, vol. 6, no. 2, pp. 22–35, 2007.
- [10] J. Yu, Y. Wang, and Y. Shen, "Noise reduction and edge detection via kernel anisotropic diffusion," *Pattern Recognition Letters*, vol. 29, no. 10, pp. 1496–1503, 2008.
- [11] T. Barbu, V. Barbu, V. Biga, and D. Coca, "A PDE variational approach to image denoising and restoration," *Nonlinear Analysis: Real World Applications*, vol. 10, no. 3, pp. 1351–1361, 2009.
- [12] A. Douiri, M. Schweiger, J. Riley, and S. R. Arridge, "Anisotropic diffusion regularization methods for diffuse optical tomography using edge prior information," *Measurement Science and Technology*, vol. 18, no. 1, pp. 87–95, 2007.
- [13] M. Basu, "Gaussian-based edge-detection methods—a survey," *IEEE Transactions on Systems, Man and Cybernetics Part C*, vol. 32, no. 3, pp. 252–260, 2002.
- [14] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Pearson, Upper Saddle River, NJ, USA, 2002.
- [15] J. F. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [16] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, vol. 80 of *Monographs in Mathematics*, Birkhäuser, Basel, Switzerland, 1984.
- [17] V. B. S. Prasath and A. Singh, "Edge detectors based anisotropic diffusion for enhancement of digital images," in *Proceedings of the 6th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '08)*, pp. 33–38, IEEE Computer Society Press, Bhubaneswar, India, December 2008.
- [18] F. Demengel and R. Temam, "Convex functions of a measure and applications," *Indiana University Mathematics Journal*, vol. 33, no. 5, pp. 673–709, 1984.
- [19] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, Fla, USA, 1992.
- [20] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland Mathematics Studies, no. 5, North-Holland, Amsterdam, The Netherlands, 1973.