RIGOROUS MATHEMATICAL INVESTIGATION OF A NONLINEAR ANISOTROPIC DIFFUSION-BASED IMAGE RESTORATION MODEL

TUDOR BARBU, ANGELO FAVINI

Abstract. A nonlinear diffusion based image denoising technique is introduced in this paper. The proposed PDE denoising and restoration scheme is based on a novel diffusivity function that uses an automatically detected conductance parameter. A robust mathematical treatment is also provided for our anisotropic diffusion model. We demonstrate that edge-stopping function model is properly chosen, explaining the mathematical reasons behind it. Also, we perform a rigorous mathematical investigation on of the existence and uniqueness of the solution of our nonlinear diffusion equation. This PDE-based noise removal approach outperforms most diffusion-based methods, producing considerably better smoothing results and providing a much better edge preservation.

1. Introduction

An efficient noise removal that preserves the essential image features, like boundaries, corners and other sharp structures, still represents a challenging task in the image processing domain [12]. The conventional image smoothing algorithms, such as averaging filter, median filter or classic Gaussian filter are capable to reduce the noise amount, but also have the disadvantage of blurring the edges [5]. For this reason, many edge-preserving denoising techniques based on Partial Differential Equations (PDEs) have been developed in the last three decades [12, 6].

The linear diffusion models represent the simplest PDE-based image denoising solutions. A 2D Gaussian smoothing process is equivalent to a linear diffusion filtering. Thus, if a degraded image is processed by convolution with a Gaussian kernel, the result represents also the solution of the heat equation [6]. The major disadvantage of the linear PDE models is their blurring effect over the image details. Linear diffusion has no localization property and could dislocate the edges when moving from finer to coarser scales.

The nonlinear diffusion based methods avoid the blurring and localization problems of the linear filters. They perform a directional diffusion, which is degenerate along the gradient direction, having the effect of denoising the image along but...
not across the edges. Various nonlinear diffusion-based noise removal techniques have been elaborated since the early work of Perona and Malik in 1987 [13]. They developed an anisotropic diffusion framework for image denoising and restoration, which was able to smooth the noisy image while preserving its edges, by encouraging the diffusion within image regions and prohibiting it across strong boundaries. Two diffusivity function variants were considered for this model [13]. There is a lot of literature based on the denoising scheme proposed by Perona and Malik, numerous nonlinear diffusion techniques derived from this influential algorithm being constructed in the last 25 years [18]. Many papers consider mathematical investigations, numerical implementations and possible applications of Perona-Malik model. Its stability has been extensively studied in the last decades [18].

Also, in the last decades there have been developed a lot of denoising approaches based on Total Variation (TV) regularization [12, 15]. The total variation principle was introduced by Rudin, Osher and Fetamì in 1992 [15]. Their variational filtering technique is based on the minimization of the TV norm. While TV denoising is remarkably effective at simultaneously preserving the edges whilst smoothing away noise in flat regions, it also suffers from the staircasing effect and its corresponding Euler-Lagrange equation is highly nonlinear and difficult to compute. For this reason, in the last two decades there have been proposed many improved versions of the TV denoising model [12], such as contrast invariant TV-$L_1$ model [10], Adaptive TV de-noising [11], spatially adaptive TV [14], anisotropic HDTV regularizer [7], TV models with ADMM algorithms [17] and TV minimization with Split Bregman [3].

We have conducted a large amount of research in the PDE-based image denoising and restoration domain in the last years, too. Several PDE variational [1] and nonlinear diffusion based techniques [2] have been developed by us in the last years. In this paper we consider a nonlinear anisotropic diffusion scheme for image restoration. The proposed model, based on a novel edge-stopping function and a conductance parameter depending on the current state of the image, is detailed in the next section. A robust mathematical treatment of the proposed diffusion scheme, representing the main contribution of this article, is provided in the third section. The image noise removal results and the performed method comparison are described in the fourth section. This paper ends with a conclusions section and a list of references.

2. Robust anisotropic diffusion based noise reduction technique

We consider an edge-preserving image noise removal PDE model using the nonlinear anisotropic diffusion. The proposed diffusion-based image noise reduction algorithm is given by the following parabolic equation:

\[
\frac{\partial u}{\partial t} = \text{div}(\psi_{K(u)}(\|\nabla u\|^2\nabla u))
\]

\[
u = u_0, \quad (x, y) \in \Omega
\]

\[
\nabla u \cdot \nu = 0, \quad \text{on} \ (0, T) \times \partial \Omega
\]

where $u_0$ is the initial noisy image, its domain is $\Omega \subset \mathbb{R}^2$ and $\nu$ is the normal to $\partial \Omega$. The nonlinear diffusion model provided by (2.1) is based on the following diffusivity
(edge-stopping) function, $ψ_{K(u)} : [0, ∞) \rightarrow [0, ∞)$:

$$ψ_{K(u)}(s^2) = \begin{cases} \alpha \sqrt{\frac{K(u)}{s^2 + \eta}}, & \text{if } s > 0, \\ 1, & \text{if } s = 0, \end{cases}$$

(2.2)

where $\alpha, \beta \in [0.5, 0.8]$ and $\eta \in [0.5, 1)$.

As one can see in (2.2), the modeled edge-stopping function is based on a conductance diffusivity depending on the state of the image $u$ at time $t$. Conductance parameter is very important for the diffusion process. When the gradient magnitude exceeds its value, the corresponding edge is enhanced [6, 13, 18]. Some algorithms, including the Perona–Malik filter, use a fixed value [13, 18]. Another solution is to make this parameter a function of time. One may use a high value at the beginning, then it is reduced gradually, as the image is smoothed [18].

Other approaches detect automatically the conductance diffusivity as a function of the current state of the processed image. Various noise estimation methods are used for conductance parameter detection [16]. We also consider an automatic computation of this parameter, based on the image noise estimation at each iteration. Some statistics are utilized by the proposed conductance parameter model that is expressed as the following function:

$$K(u) = \|u\|_F \frac{\text{median}(u)}{\varepsilon \cdot n(u)},$$

(2.3)

where $\varepsilon \in (0, 1]$, $\|u\|_F$ is the Frobenius norm of image $u$, median($u$) represents its median value and $n(u)$ is the number of its pixels.

The proposed diffusivity function $ψ_{K(u)}$ is properly chosen. In the next section, where a mathematical investigation of the developed model is provided, we demonstrate that $ψ_{K(u)}$ satisfies the main conditions related to any edge-stopping function. The problem of existence and uniqueness of the solution of our anisotropic diffusion model is also investigated in the third section, where we prove that this equation has a unique weak solution in some certain cases.

A robust numerical approximation scheme is then computed for this continuous mathematical model. Thus, the equation (2.1) is discretized by using a 4-nearest-neighbours discretization of the Laplacian operator, $Δu$. From (2.1), we obtain

$$\frac{∂u}{∂t} = \text{div}(ψ_{K(u)}(|\nabla u|^2)\nabla u) \Rightarrow u(x, y, t + 1) - u(x, y, t) \cong \text{div}(ψ_{K(u)}(|\nabla u|^2)Δu),$$

which leads to the approximating scheme

$$u^{t+1} = u^t + \lambda \sum_{q \in N(p)} ψ_{K(u)}(|\nabla u^{p,q}(t)|) \nabla u^{p,q}(t),$$

(2.4)

where $λ \in (0, 1)$, $N(p)$ is the set of pixels representing the 4-neighborhood of the pixel $p = (x, y)$ ($x$ and $y$ representing coordinates) and the image gradient magnitude in a particular direction at iteration $t$ is computed as follows:

$$\nabla u^{p,q}(t) = u(q, t) - u(p, t).$$

(2.5)

The restoration algorithm given by (2.4) is applied on the current image for $t = 0, 1, \ldots, N$, where $N$ is the maximum number of iterations. Our noise removal iterative approach converges quite fast to the desired solution. More about the convergence of this finite difference scheme is discussed in the next section. It produces the smoothed image $u^N$ from the degraded image $u^0 = u_0$ in a relatively low number of steps, therefore the $N$ value has to be quite low. The experiments performed by using this iterative scheme are described in the fourth section.
3. Mathematical treatment of the anisotropic diffusion model

In this section we provide a mathematical treatment of the proposed nonlinear anisotropic diffusion model. First, we have to demonstrate that \( \psi_{K(u)} \) is properly modeled, satisfying the main properties of an efficient edge-stopping function \[6, 13, 17\]. Obviously, we have \( \psi_{K(u)}(0) = 1 \). Also, the function is always positive, because \( \alpha \cdot \sqrt{\frac{K(u)}{\beta s^2 + \eta}} > 0, \forall s \in \mathbb{R} \). The function \( \psi_{K(u)}(s^2) \) is monotonically decreasing, because

\[
\psi_{K(u)}(s_1^2) = \alpha \sqrt{\frac{K(u)}{\beta \cdot s_1^2 + \eta}} \leq \alpha \sqrt{\frac{K(u)}{\beta s_2^2 + \eta}} \leq \psi_{K(u)}(s_2^2)
\]

for all \( s_1 \geq s_2 \). We also have \( \lim_{s \to \infty} \psi_{K(u)}(s^2) = 0 \).

Besides these conditions, our diffusivity function satisfies another important one. If one considers the flux function, defined as \( \phi(s) = s \cdot \psi_{K(u)}(s^2) \), the process of enhancing the image and sharpening its edges depends on the sign of its derivative, \( \phi'(s) > 0 \). So, if the derivative of a flux function of a diffusion model is positive \( \phi'(s) > 0 \), then the respective model represents a forward parabolic equation. Otherwise, for \( \phi'(s) < 0 \), that nonlinear diffusion model is a backward parabolic equation \[8\]. In our case, the derivative of the flux function is computed as:

\[
\phi'(s) = \psi_{K(u)}(s^2) + 2s^2 \psi'_{K(u)}(s^2)
\]

which leads to

\[
\phi'(s) = \frac{\alpha \sqrt{K(u)}}{\beta s^2 + \eta} - \frac{\alpha \sqrt{K(u)} s^2}{\beta s^2 + \eta} \cdot \frac{\beta}{\sqrt{\beta s^2 + \eta}}.
\]

Therefore,

\[
\phi'(s) = \frac{\alpha \sqrt{K(u)}}{(\beta s^2 + \eta)^{3/2}} [\beta s^2 + \eta - \beta s^2] = \frac{\alpha \gamma \sqrt{K(u)}}{(\beta s^2 + \eta)^{3/2}}.
\]

Since \( \frac{\alpha \gamma \sqrt{K(u)}}{(\beta s^2 + \eta)^{3/2}} > 0 \), we obtain \( \phi'(s) > 0 \) for any \( s \), which means our PDE denoising model is a forward parabolic equation that is stable and it is quite likely to have a solution.

The existence and uniqueness of the solution of our proposed diffusion model requires a robust mathematical investigation. It should be said that, in general, the problem \[2.1\] is ill-posed. It does not have a classical solution but has a solution in weak sense that is in sense of distributions. One can prove the existence and uniqueness of a weak solution in a certain case, related to some values of the parameters of this model. Thus, we demonstrate that our anisotropic diffusion model converges if \( \gamma = \alpha^2 \). Let us consider the following modification of the function \( \psi_{K(u)} \):

\[
\psi_{K(u)}(s^2) = \begin{cases} \frac{\alpha \sqrt{K(u)}}{\beta s^2 + \eta}, & s \in (0, M], \\ \frac{\alpha}{\sqrt{s}}, & \text{if } s = 0, \end{cases}
\]

where \( M > 0 \) is arbitrarily large but fixed. The function \( K \) given by \[2.3\] is Lipschitz and positive, that is, \( |K(u) - K(v)| \leq \ell \cdot |u - v| \) and \( K(u) \geq \rho \), \( \forall u \). By
We consider the set unique weak solution $u$ to problem

\begin{equation}
\int \frac{\partial}{\partial t} u(t, x, y) \varphi(x, y) \, dx \, dy = - \int \Omega \psi_{K(u)}(|\nabla u(t, x, y)|^2) \nabla u(t, x, y) \cdot \nabla \varphi(x, y) \, dx \, dy
\end{equation}

\begin{equation}
u(0, x, y) = u_0(x, y), \quad \forall (x, y) \in \Omega.
\end{equation}

Here $L^2(\Omega)$ is the space of all Lebesgue square integrable functions on $\Omega$ and the Sobolev space $H^1(\Omega) = \{u \in L^2(\Omega), \frac{\partial u}{\partial x_i} \in L^2(\Omega), i = 1, 2\}$, where $\frac{\partial u}{\partial x_i}$ is taken in the sense of distributions.

We have denoted by $(H^1(\Omega))'$ the dual of $H^1(\Omega)$ and by $C([0, 1]; L^2(\Omega))$ the space of continuous functions $u : [0, 1] \rightarrow L^2(\Omega)$. By $L^2(0, T; H^1(\Omega))$, respectively $L^2(0, T; (H^1(\Omega))')$, we denote the space of measurable functions $u : (0, T) \rightarrow H^1(\Omega)$ (respectively, $u : (0, T) \rightarrow (H^1(\Omega))'$) such that

\begin{equation}
\int_0^T \|u(t)\|_{H^1(\Omega)} \, dt < \infty, \quad \text{respectively} \quad \int_0^T \|u(t)\|_{(H^1(\Omega))'}^2 \, dt < \infty.
\end{equation}

**Proposition 3.1.** Assume that (2.2) holds. Then, for each $u_0 \in L^2(\Omega)$ there is a unique weak solution $u$ to problem (2.1). Moreover, if $u_0 \geq 0$, then $u \geq 0$.

**Proof.** We consider the set

\begin{equation}
\chi = \{u \in C([0, 1]; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega)); \|u\|_{L^2(0, T; H^1(\Omega))} \leq R\}
\end{equation}

and fix $v \in \chi$. Consider the Cauchy problem

\begin{equation}
\frac{\partial u}{\partial t} = \text{div}(\psi_{K(v)}(\nabla u)^2 \nabla u), \quad t \in (0, T), \quad (x, y) \in \Omega,
\end{equation}

with $\nabla u \cdot v = 0$ on $(0, T) \times \partial \Omega, u(0) = u_0$ in $\Omega$. Equivalently, we have $\frac{\partial u}{\partial t} = A_v(t)u, u(0) = u_0$, with $A_v(t) : H^1(\Omega) \rightarrow (H^1(\Omega))'$:

\begin{equation}
\langle A_v(t)u, \varphi \rangle_{H^1(\Omega)} = \int \Omega \psi_{K(v)}(|\nabla u|^2) \nabla u \cdot \nabla \varphi \, dx \, dy, \quad \forall \varphi \in H^1(\Omega), \quad t \in (0, T).
\end{equation}

By a little computation involving (3.4), it follows that the operator $A_v(t)$ is, for each $t$, monotone and demicontinuous from $H^1(\Omega)$ to $(H^1(\Omega))'$. In other words, $\langle A_v(u) - A_v(\bar{u}), u - \bar{u} \rangle \geq 0, \forall u, \bar{u} \in H^1(\Omega)$ and $u \rightarrow A_v(u)$ is strongly-weakly continuous from $H^1(\Omega)$ to $(H^1(\Omega))'$. Moreover, for some $C, \alpha > 0$, we have

\begin{equation}
\|A_v(t)u\|_{(H^1(\Omega))'} \leq C\|u\|_{H^1(\Omega)}, \quad \forall u \in H^1(\Omega),
\end{equation}

\begin{equation}
\langle A_v(t)u, u \rangle_{H^1(\Omega)} \geq \alpha\|u\|_{H^1(\Omega)}^2, \quad \forall u \in H^1(\Omega).
\end{equation}

Then, according to a well-known result due to Lions [9], for each $v \in \chi$, problem (3.7) has a unique weak solution $u = \Phi(v)$. Now, it suffices to show that $\Phi$ is a contraction on $\chi$ and leaves invariant this set. A little calculation involving (3.5) and the monoticity of the mapping $r \rightarrow \psi_{K(u)}(|r|^2)r$ shows that, for some $\alpha_0 > 0$,

\begin{equation}
\frac{1}{2} \frac{\partial}{\partial t} \|u(t) - \bar{u}(t)\|_{L^2(\Omega)}^2 + \alpha_0 \int \Omega \left|\nabla \left(u(t, x, y) - \bar{u}(t, x, y)\right)\right|^2 \, dx \, dy
\end{equation}

\begin{equation}
\leq \int \Omega \left|\nabla \left(u(t, x, y) - u(t, x, y)\right)\right| K(v(t, x, y)) - K(\bar{v}(t, x, y)) \, dx \, dy,
\end{equation}

where $K$ is a positive compact operator.
for \( t \in (0, T) \). This yields
\[
\|u(t) - \bar{u}(t)\|_{L^2(\Omega)}^2 + \frac{\alpha_0}{2} \int_0^t \int_{\Omega} |\nabla (u(t, x, y) - \bar{u}(t, x, y))|^2 \, dx \, dy \, dt
\leq C \int_0^t \int_{\Omega} |v(t, x, y) - \bar{v}(t, x, y)|^2 \, dx \, dy \, dt,
\]
(3.10)
where \( u = \Phi(v) \), \( \bar{u} = \Phi(\bar{v}) \). If in \( \chi \) we will consider the metric defined by the distance \( d(u, \bar{u}) = \sup_{0 \leq t \leq T} \|u(t) - \bar{u}(t)\|_{L^2(\Omega)} \) with \( \gamma_0 \) suitably chosen, it follows that \( d(\Phi(v), \Phi(\bar{v})) \leq \rho d(v, \bar{v}), \forall v, \bar{v} \in X \), for some \( 0 < \rho < 1 \). Moreover, it follows by a similar calculation that, for a suitably chosen \( R \), \( \Phi \) leaves invariant the set \( \chi \). Then, by Banach’s fixed point theorem, it follows the existence and uniqueness in \( (2.1) \). Moreover, if \( u_0 \geq 0 \), then, taking \( \varphi = u^- \) in \( (3.5) \), we see after some calculation that \( u^- = 0 \) and so \( u \geq 0 \). This completes the proof of Proposition 3.1. We also have several remarks.

1. By replacing function \( (2.1) \) by \( (3.4) \), we have apparently modified the original model. However, since in specific denoising or restoring applications the magnitude of the gradient does not exceed a certain value (even for sharp edges), this choice of \( \psi_{K(u)} \) is reasonable for \( M \) sufficiently large.

2. By Proposition 3.1 and its proof, it follows that the solution \( u \) to \( (2.1) \) can be obtained iteratively as \( u = \lim_{n \to \infty} u_n \), where \( u_n \) represents the weak solution to the problem \( (2.1) \). Moreover, this implies that the solution \( u \) to \( (2.1) \) can be obtained as limit of the finite difference scheme mentioned in the previous section:
\[
u(t+1) = u(t) + \text{div} (\psi_{K(u)}(|\nabla u|^2) \nabla u) \quad \text{in} \; \Omega, \quad u(t) \cdot \nu = 0 \; \text{on} \; \partial \Omega.
\]
(3.11)

3. We may rewrite equation \( (2.1) \) as
\[
\frac{\partial}{\partial t} \sqrt{K(u)} = \frac{1}{2} \text{div} (g_0(|\nabla u|^2) \nabla u) + \frac{1}{4} \frac{K'(u)}{K(u)} g_0(|\nabla u|^2)(|\nabla u|^2) \quad \text{in} \; (0, T) \times \Omega,
\]
(3.12)
with \( u(0) = u_0 \in \Omega, \nabla u \cdot \nu = 0 \) on \( (0, T) \times \partial \Omega \), where \( g_0(s) = \frac{s}{\sqrt{1+s^2}} \) for \( s > 0 \). If we neglect the low order term, then we get the equation
\[
\frac{\partial}{\partial t} \sqrt{K(u)} = \frac{1}{2} \text{div} (g_0(|\nabla u|^2) \nabla u) \quad \text{in} \; (0, T) \times \Omega
\]
(3.13)
with the Neumann boundary condition. This is a nonlinear parabolic equation of the form studied in [4] and the methods developed there can be used to obtain existence in \( (2.1) \) under more general conditions on \( K(u) \) (for instance, for \( K(u) > 0 \)). It should be said that, for \( K(u) = \text{constant} \), problem \( (2.1) \) reduces to the bounded variation flow model and it is well posed in the space of functions with bounded variation.

4. Experiments

The described anisotropic diffusion-based denoising technique has been tested on hundreds images affected by various levels of Gaussian noise, satisfactory results being achieved. The following parameters of our PDE model have provided the optimal smoothing results: \( \alpha = 0.7, \beta = 0.65, \eta = 0.5, \varepsilon = 0.3, \lambda = 0.33 \) and \( N = 15 \). One can see that \( \eta \approx \alpha^2 \), therefore the nonlinear diffusion scheme has a unique solution and it converges fast to it, the \( N \) value being quite low.
A method comparison has also been performed. The proposed anisotropic diffusion approach outperforms many other denoising methods, obtaining better noise reduction results and converging considerably faster than some PDE-based algorithms, including the Perona-Malik scheme [13] and the Total Variation (TV) techniques [10, 11, 15]. It also provides a much better smoothing than the non-PDE based filtering methods [5].

In Figure 1 there are displayed: (a) the original [512 × 512] Peppers image; (b) the image corrupted with Gaussian noise characterized by parameters $\mu = 0.21$ and $\text{var} = 0.02$; (c) the image processed by our AD (anisotropic diffusion) technique; (d) the image filtered by the Perona-Malik algorithm; (e) the image denoised by a TV regularization scheme; (f)–(i) denoising results of $[3 \times 3]$ 2D Gaussian, average, median and Wiener kernels [5]. It is obvious from these displays that our approach produces the best edge-preserving restoration results.

![Figure 1. Method comparison: image processed with several filtering techniques](image)

The performance of our denoising method has been assessed by using the norm of the error image measure, computed as $\sqrt{\sum_{x=1}^{X} \sum_{y=1}^{Y} [u(x,y) - u_0(x,y)]^2}$. The
Table 1. Norm-of-the-error measure values for various denoising algorithms

<table>
<thead>
<tr>
<th></th>
<th>Our AD</th>
<th>P-M</th>
<th>TV</th>
<th>Gaussian</th>
<th>Average</th>
<th>Median</th>
<th>Wiener</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>5.15 × 10^3</td>
<td>6.1 × 10^3</td>
<td>5.8 × 10^3</td>
<td>7.3 × 10^3</td>
<td>6.4 × 10^3</td>
<td>6 × 10^3</td>
<td>5.9 × 10^3</td>
</tr>
</tbody>
</table>

NE image values corresponding to the experiments described in Figure 1 are registered in Table 1. As one can see in this table, our AD approach corresponds to the lowest NE value that indicates the best denoising.

**Conclusion.** We have proposed a nonlinear anisotropic diffusion based model for image noise reduction in this paper. Our robust PDE technique performs not only an efficient image denoising, but also an enhancement of the image boundaries.

A novel edge-stopping function and a conductance parameter modeled as a function of processed image are constructed, as well as an efficient numerical approximation iterative algorithm, but the major contribution of this article is the rigorous mathematical treatment of the developed PDE-based model. First, we have proved that the modeled diffusivity function is properly chosen, satisfying the required properties. Then, we have performed a mathematical demonstration of the existence and uniqueness of the solution of this forward parabolic equation. We demonstrate that this equation has a unique weak solution in some certain cases.

The performed restoration tests and the method comparison provided satisfactory results. The developed nonlinear diffusion approach outperforms both the PDE-based algorithms, like the Perona–Malik scheme or TV model, and the conventional filtering techniques. Also, given its robust edge-preserving character, our technique described here can be successfully used for solving edge detection and image object detection tasks.

**Acknowledgments.** This research has been supported by the project PN-II-PCE-20113-0027, financed by the Romanian Minister of Education and Technology. We acknowledge also the research support of the Institute of Computer Science of the Romanian Academy, Iași, Romania.

**References**


TUDOR BARBU  
INSTITUTE OF COMPUTER SCIENCE OF THE ROMANIAN ACADEMY, IASI, ROMANIA  
E-mail address: tudor.barbu@iit.academiaromana-is.ro

ANGELO FAVINI  
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BOLOGNA, BOLOGNA, ITALY  
E-mail address: angelo.favini@unibo.it