INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 18, Number 1, Pages 38–61

MULTI-SCALE NON-STANDARD FOURTH-ORDER PDE IN IMAGE DENOISING AND ITS FIXED POINT ALGORITHM

ANIS THELJANI

Abstract. We consider a class of nonstandard high-order PDEs models, based on the $(p(\cdot), q(\cdot))$-Kirchhoff operator with variable exponents for the image denoising problem. We theoretically analyse the proposed non-linear model. Then, we use linearization method based on a fixed-point iterative technique and we also prove the convergence of the iterative process. The model has a multiscale character which follows from an adaptive selection of the exponents $p(\cdot)$ and $q(\cdot)$. The latter task helps to capture, highlight and correlate major features in the images and optimize the smoothing effect. We use Morley finite-elements for the numerical resolution of the proposed model and we give several numerical examples and comparisons with different methods.

Key words. High-order PDEs, fixed point method, anisotropic diffusion, finite elements, image restoration, inverse problems

1. Introduction

Image restoration is a fundamental task in image processing and it arises in diverse fields such as geophysics, optics, medical imaging[33, 35, 37]. It is a classical inverse problem which aims at reconstructing an image $u : \Omega \to \mathbb{R}$ from an observed one $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ that is degraded and contaminated by noise. The degradation model that we consider is the following:

$$f = u + \eta, \tag{1}$$

where η is Gaussian noise. Estimating u from the model (1) is an ill posed inverse problem where a prior image model $\mathcal{R}(u)$ is required in order to successfully estimate u from the observations f. To incorporate a prior image model $\mathcal{R}(u)$ into (1), variational approach is usually used and it consists in solving a minimization problem that have the following form:

$$\min_u \left\{ \mathcal{J}(u) := \mathcal{R}(u) + \tfrac{\lambda_0}{2} \|u - f\|^2_{L^2(\Omega)} \right\}. \tag{2}$$

The prior $\mathcal{R}(\cdot)$ in the energy $\mathcal{J}(\cdot)$ have a regularization effect and usually contains information about the image derivatives to reduce the noise that is considered as high oscillations. The second part of the energy $\mathcal{J}(\cdot)$ is the fitting term, λ_0 is a positive regularization parameter which controls the trade-off between the two terms.

A main issue in image denoising is how to choose the "best" regularization term $\mathcal{R}(\cdot)$ that can selectively smooth a noisy image without losing significant features such as edges and thin structures. Various regularizers based on first- or/and second-order derivatives have been used [14, 7, 10, 40, 25]. In [33], the authors proposed to use the well-known total variation (TV) regularizer $\mathcal{R}(u) = TV(u)$ where

$$TV(u) := \int_\Omega |Du| = \sup \left\{ \int_\Omega u \mathrm{div}\varphi dx | \varphi \in C^2_c(\Omega, \mathbb{R}^2),\ |\varphi| \le 1 \right\},$$

Received by the editors October 10, 2020.
2000 *Mathematics Subject Classification.* 65M32, 47J25, 65M50, 65M22, 94A08, 35G30, 35Q68, 65J15.

which produces a piecewise constant restored images. However, TV also produces staircase effects which is undesirable. This shortcoming gave rise to a class of a combined first- and second-order derivatives as regularizer that in general damp the noise faster and diminish the staircase effect. There have been many efforts to improve the robustness and to reduce the staircasing effects of TV using the high-order TV and total generalized variation (TGV) regularizer [11, 14]. Most of the high-order models aim to extend the works in [12] (see also, e.g., [31, 38, 39, 43]) which uses straightforward convex combinations of first- and second- derivatives. They are generally written in following form:

$$\int_\Omega G_1(\nabla u)\,dx + \int_\Omega G_2(\nabla^2 u)\,dx + \tfrac{\lambda_0}{2}\|u-f\|^2_{L^2(\Omega)}, \tag{3}$$

where $G_1(\cdot)$ and $G_2(\cdot)$ are given functions. In [41], a high-order total variation model, called $TV - TV^2$, was proposed and it consists in minimizing the following energy:

$$\alpha TV(u) + \beta TV^2(u) + \tfrac{\lambda_0}{2}||u-f||^2_{L^2(\Omega)}, \tag{4}$$

where α and β are non-negative regularization parameters chosen empirically, $TV(u)$ and $TV^2(u)$ are the total variations of u and ∇u, respectively.

Various variations of high-order models that are based on the above two energies forms were proposed [44, 48, 26, 28]. Most of these models gave rise to a second- or high-order non-linear PDEs that only consider nonlinear diffusion to denoise the image. However, nonlinear diffusion is not always the best choice for homogeneous regions, i.e. no edges but only some noise. In these regions, using linear diffusion is more appropriate as it damps noise better than nonlinear diffusion. Ideally, there should be a compromise between linear diffusion PDEs which are more interesting and effective in homogeneous regions, and nonlinear diffusion PDEs that are more powerful in regions containing edges and details.

Another class of approaches, known as nonstandard PDEs with $p(\cdot)$-growth conditions were also considered in several works (see e.g., [5, 46, 24, 36, 32]). In these approaches, the regularizer takes the form of

$$\mathcal{R}(u) = \int_\Omega |\nabla u|^{p(x)} dx,$$

where $1 \leq p \leq 2$. The two extreme values of the exponent $p = 1, 2$ in the regularization term lead to nonlinear (selective) diffusion and linear (isotropic) diffusion equations. In fact, the total variation model is obtained for $p = 1$ and which leads to a nonlinear diffusion PDEs where the diffusion is guided by the term $\frac{1}{|\nabla u|}$. Thus the diffusion will be selective and inverse proportional to $|\nabla u|$, i.e. for edges where $|\nabla u|$ is high, the diffusion will be enabled in order to keep edges, whereas for the homogeneous regions where $|\nabla u|$ is small, the diffusion will be strong and the model behaves similarly to a Laplace smoothness operator. For $p = 2$, the model leads to a PDE that uses the Laplace $\Delta\cdot$ as diffusion operator. The latter has an isotropic and linear diffusion property that can't distinguish between edges and homogeneous regions.

In these nonstandard regularizations, a compromise between fast/slow diffusion is made by varying $p(\cdot)$ according to the local scales. The linear diffusion is encouraged away from the edges of the image and a nonlinear correction is enforced near these singularities (see [27, 28, 26, 1, 8, 13]). To incorporate the singularity information into the $p(\cdot)$, the authors in [8] used a variable exponent $p(\cdot)$ ranging from 1 to 2 by taking $p(|\nabla u|)$ where $p(\cdot)$ is a monotone decreasing function such that:

$$\begin{cases} \lim\limits_{r\to 0} p(r) = 2, \\ \lim\limits_{r\to\infty} p(r) = 1. \end{cases}$$

However, by this choice of $p(\cdot)$ is a priori and essentially lacks a practical selection criterium which necessarily should be linked to a tight location of the singularities set of the image. Moreover, the values of $p(\cdot)$ only include information about first-order discontinues of u, however, the second-order discontinuities, i.e. discontinuities of ∇u are also significant and interesting.

In this article, we consider a nonstandard high-order nonlinear diffusion equation. Our motivation in this choice relies on two main reasons. First, we use high-order diffusion models which, in general, perform better than second-order ones in image denoising. Second, the diffusion is driven by a variable exponents $p(\cdot), q(\cdot)$ of the first and second-order derivative regularizes and which are ranging between 1 and 2. Most importantly, we provide a practical tool for the selection of $p(\cdot), q(\cdot)$ possessing three nice aspects: being adaptive, local and linked with a tight location of the singularities set. More precisely, we consider the following minimization problem:

$$\min_u \left\{ \mathcal{J}(u) = \int_\Omega \tfrac{1}{p(x)} |\Delta u|^{p(x)} dx + \int_\Omega \tfrac{1}{p(x)} |\nabla u|^{q(x)} dx + \tfrac{\lambda_0}{2} \|u - f\|^2_{L^2(\Omega)} \right\}, \tag{5}$$

where the functions $p(\cdot)$ and $q(\cdot)$ are defined on Ω and satisfies $1 < p(\cdot), q(\cdot) \leq 2$. Then, we make an adaptive choice of the exponents $p(\cdot)$ and $q(\cdot)$ allowing slow diffusion near edges, and fast diffusion in the homogeneous regions. The selection is performed at the discrete level with the help of suitable a posteriori error indicators, i.e., no guess on the continuous solution u is required but only its finite-element approximation. The error indicators are computed based on information that comes from a smoothed structure tensors (SST) of the image and its gradient. The SST has been extensively used in nonlinear anisotropic PDEs and usually allows to capture all geometric features in the image at different scales, i.e. flat edges, corners, smooth regions, thin structures etc, see [47, 49]. Thus, the new model will have anisotropic nonlinear diffusion property and acts in a multiscale level by adopting the diffusion dependant on the region. This will help to effectively reduce noise and keep the important feature sharp.

Compared to the other mentioned high-order models, the new one combines the effectiveness of both linear and nonlinear diffusion models by varying the exponents between $(1, 2]$. Moreover, it is able to use the geometric information extracted from the SST and its anisotropy in guiding the diffusion and the smoothness process to sharpen edges and corners. Thus, it is inherently dependent on the structure tensor estimation, which therefore determines the performance of the method. In addition, we use a mesh-adaptation technique allowing to get the tight location of the singularities and permits both the refinement (near the edges) and the coarsening of the grid (in homogeneous area) in order to best fit the geometry of the image and to make the method considerably fast. This may be also seen as edge enhancing via the mesh adaptation in order to get sharp edges as much as possible.

The rest of the paper is organized as follows: In section 2, we fix notations and present preliminary results. In Section 3, we prove that our model admits a unique solution. Then, we regularize the proposed model and prove the convergence of the regularized solutions to a solution of the original problem. In section 4, we construct a sequence of linearized problems such that the sequence of their solutions converges in H^2-norm to the solution of our model. In Section 5, we detail the discretization of the equations and the adaptive strategy for the selection of the exponent $p(\cdot)$.

Finally, we illustrate in Section 6 the efficiency and robustness of the proposed method by solving a wide range of examples.

2. Preliminaries

We give some definitions and basic properties of the generalized Lebesgue and Sobolev spaces. Interested readers may refer to [17, 19] for more details. For a variable exponent $p(\cdot) \in C(\overline{\Omega})$ such that $1 < p^- \leq p(\cdot) \leq p^+ \leq 2$, we define the variable exponent Lebesgue space $L^{p(\cdot)}(\Omega)$ as follows:

$$L^{p(\cdot)}(\Omega) = \left\{ u : \Omega \to \mathbb{R}; u \text{ is measurable and } \int_\Omega |u|^{p(x)} dx < \infty \right\},$$

which is equipped with the following Luxemburg norm:

$$\|u\|_{L^{p(\cdot)}} = \inf \left\{ \nu > 0 : \int_\Omega \left| \frac{u(x)}{\nu} \right|^{p(x)} dx \leq 1 \right\}.$$

Similarly, the Sobolev space with variable exponent $W^{k,p(\cdot)}(\Omega)$ is defined as:

$$W^{k,p(\cdot)}(\Omega) = \left\{ u \in L^{p(\cdot)}(\Omega) : D^\xi u \in L^{p(\cdot)}(\Omega), |\xi| \leq k \right\},$$

where $D^\xi u = \frac{\partial^{|\xi|}}{\partial x_1^{\alpha_1} \partial x_2^{\xi_2} \ldots \partial x_N^{\xi_N}} u$ with $\xi = (\xi_1, \ldots, \xi_N) \in \mathbb{N}^N$ is a multi-index and $|\xi| = \sum_{i=1}^N \xi_i$. The space $W^{k,p(\cdot)}(\Omega)$, is equipped with the norm:

$$\|u\|_{k,p(\cdot)} := \sum_{|\xi| \leq k} ||D^\xi u||_{L^{p(\cdot)}}.$$

We recall that both $(L^{p(\cdot)}(\Omega), || \cdot ||_{L^{p(\cdot)}})$ and $(W^{k,p(\cdot)}(\Omega), || \cdot ||_{k,p(\cdot)})$ are separable, reflexive and uniformly convex Banach spaces [21]. For more details, we refer the reader to [19, 20]. The norm $\|u\|_{2,p(\cdot)}$ and $\|u\|_{1,q(\cdot)}$ are respectively equivalent to the following norms

$$\|u\|_p = \inf \left\{ \varrho > 0 : \int_\Omega \left| \frac{\Delta u(x)}{\varrho} \right|^{p(x)} \mathrm{d}x + \int_\Omega \left| \frac{\nabla u(x)}{\varrho} \right|^{p(x)} \mathrm{d}x + \int_\Omega \left| \frac{u(x)}{\varrho} \right|^{p(x)} \mathrm{d}x \leq 1 \right\},$$

$$\|u\|_q = \inf \left\{ \varrho > 0 : \int_\Omega \left| \frac{\nabla u(x)}{\varrho} \right|^{q(x)} \mathrm{d}x + \int_\Omega \left| \frac{u(x)}{\varrho} \right|^{q(x)} \mathrm{d}x \leq 1 \right\},$$

in the spaces $W^{2,p(\cdot)}(\Omega)$ and $W^{1,q(\cdot)}(\Omega)$ (for more details see [6]). Moreover, $(W^{2,p(\cdot)}(\Omega); \|\cdot\|_p)$ and $(W^{1,q(\cdot)}(\Omega); \|\cdot\|_q)$ are Banach, separable and reflexive spaces.

In the sequel. we consider the following space

$$X = \{u \in W^{2,p(\cdot)} \cap W^{1,q(\cdot)}(\Omega) \,|\, u = \frac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0\},$$

equipped with the norm

$$\|u\|_X = (\|u\|_p^2 + \|u\|_q^2)^{1/2}.$$

Then, $(X; \|\cdot\|_X)$ is Banach, separable and reflexive space.

Lemma 2.1. [22] *For all $u \in X$, we have:*

$$\begin{aligned} &\min\{\|\nabla u\|_{L^{q(\cdot)}}^{q+}, \|\nabla u\|_{L^{q(\cdot)}}^{q-}\} + \min\{\|\Delta u\|_{L^{p(\cdot)}}^{p+}, \|\Delta u\|_{L^{p(\cdot)}}^{p-}\} \\ \leq &\int_\Omega |\nabla u|^{q(x)} \, dx + \int_\Omega |\Delta u|^{p(x)} \, dx \\ \leq &\max\{\|\nabla u\|_{L^{q(\cdot)}}^{q+}, \|\nabla u\|_{L^{q(\cdot)}}^{qp-}\} + \max\{\|\Delta u\|_{L^{p(\cdot)}}^{p+}, \|\Delta u\|_{L^{p(\cdot)}}^{p-}\}. \end{aligned}$$

Lemma 2.2 ([18]). **(Poincaré's inequality).** *If Ω is a Lipschitz domain, then, there exists a constant $C > 0$ such that:*

$$\|u\|_{L^{q(\cdot)}} \leq C\|\nabla u\|_{L^{q(\cdot)}}, \quad \forall u \in W_0^{1,q(\cdot)}(\Omega).$$

3. Existence of weak solution

In the sequel, we will establish the existence and uniqueness of a solution for our model in the space X.

Proposition 1. *For fixed $f \in L^2(\Omega)$, the minimization problem (5) admits a unique solution u in X. Moreover, the solution u fulfils*

$$\begin{cases} \Delta\left(|\Delta u|^{p(x)-2}\Delta u\right) - \nabla.(|\nabla u|^{q(x)-2}\nabla u) + \lambda_0 u = \lambda_0 f, & \text{in } \Omega, \\ u = \frac{\partial u}{\partial n} = 0, & \text{on } \partial\Omega. \end{cases} \tag{6}$$

Proof. The energy $\mathcal{J}(\cdot)$ is weakly lower semi-continuous in X. Now, let $(u_n)_{n\in\mathbb{N}}$ be a minimizing sequence in X of $\mathcal{J}(\cdot)$, then there exists $C > 0$ such that

$$\mathcal{J}(u_n) \leq C. \tag{7}$$

Thus, $\mathcal{R}(u_n) \leq C$ and hence $\|\Delta u_n\|_{L^{p(\cdot)}}$ and $\|\nabla u_n\|_{L^{q(\cdot)}}$ are uniformly bounded. Then, using Proposition 2.1 and Poincare's inequality, we obtain that $(u_n)_{n\in\mathbb{N}}$ is uniformly bounded in X which means that there exists a subsequence, still denoted $(u_n)_{n\in\mathbb{N}}$, such that $u_n \underset{n\to\infty}{\rightharpoonup} u$ weakly in X and the limit u is a minimiser of $\mathcal{J}(\cdot)$. By using the lower semi-continuity of $\mathcal{J}(\cdot)$ and Fatou's Lemma we obtain that u is a minimizer for $\mathcal{J}(\cdot)$. Uniqueness follows from the strict convexity of $\mathcal{J}(\cdot)$. In addition, for $X \setminus \{0\}$ we have:

$$\begin{aligned} \langle \mathcal{J}'(u), v\rangle &= \tfrac{d}{dt}\left\{\mathcal{J}(u+tv)\right\}_{t=0} \\ &= \int_\Omega |\Delta u|^{p(x)-2}\Delta u \Delta v\, dx + \int_\Omega |\nabla u|^{q(x)-2}\nabla u\cdot\nabla v\, dx \\ &+ \lambda_0 \int_\Omega uv\, dx - \lambda_0\int_\Omega fv\, dx = 0, \qquad \forall v \in X. \end{aligned}$$

Therefore, the functional $\mathcal{J}(\cdot)$ is Gâteaux differentiable in $X \setminus \{0\}$ and its unique minimizer is a solution of the weak formulation: Find u in X such that:

$$\begin{aligned} &\int_\Omega |\Delta u|^{p(x)-2}\Delta u\Delta v\, dx + \int_\Omega |\nabla u|^{q(x)-2}\nabla u\cdot\nabla v\, dx + \lambda_0\int_\Omega uv\, dx \\ =&\lambda_0 \int_\Omega fv\, dx, \quad \forall v \in X, \end{aligned} \tag{8}$$

Moreover, integration the weak formulation (8) by part and using Green formula, we clearly get the PDEs (6).

□

3.1. Γ-convergence approximation.

The proposed model is not defined for the degenerate cases $|\nabla u| = 0$ or/and $|\Delta u| = 0$. Such singularities can be avoided by perturbing the term $|\nabla u|$, $|\Delta u|$ with a small positive constant ϵ, to obtain a new regularized model. More precisely, we consider the following regularization

$$|\Delta u| \approx \sqrt{\epsilon + |\Delta u|^2} \text{ and } |\nabla u| \approx \sqrt{\epsilon + |\nabla u|^2}.$$

Thus we have

$$\int_\Omega \tfrac{1}{p(x)}|\Delta u|^{p(x)}dx \approx \int_\Omega \tfrac{1}{p(x)}(\epsilon+|\Delta u|^2)^{\frac{p(x)}{2}}dx$$
$$\text{and } \int_\Omega \tfrac{1}{q(x)}|\nabla u|^{q(x)}dx \approx \int_\Omega \tfrac{1}{q(x)}(\epsilon+|\nabla u|^2)^{\frac{q(x)}{2}}dx.$$

Moreover, since we are dealing with non-linear PDEs that needs linearization techniques, we will use a fixed point approach that requires some ellipticity assumptions in the space $\tilde{H}_0^2(\Omega) = \{u \in W^{2,2}(\Omega) \,|\, u = \frac{\partial u}{\partial n}\big|_{\partial\Omega} = 0\}$. Thus, we add the term $\epsilon^2\int_\Omega |\Delta u|^2dx$ to original energy to guarantee the ellipticity. More precisely, we have used the approximation

$$\int_\Omega \tfrac{1}{p(x)}(\epsilon+|\Delta u|^2)^{\frac{p(x)}{2}}dx \approx \int_\Omega \left[\tfrac{1}{p(x)}(\epsilon+|\Delta u|^2)^{\frac{p(x)}{2}}dx + \epsilon^2|\Delta u|^2\right]dx$$

for small $\epsilon > 0$. Therefore, the minimization problem (5) is approximated by the following optimization problem

(9)
$$\min_{u\in\tilde{H}_0^2(\Omega)}\left\{\mathcal{J}_\epsilon(u) = \int_\Omega \tfrac{1}{p(x)}K_{\epsilon,p}(|\Delta u|^2)dx + \int_\Omega \tfrac{1}{q(x)}K_{\epsilon,q}(|\nabla u|^2)dx + \tfrac{\lambda_0}{2}\|u-f\|^2_{L^2(\Omega)}\right\},$$

where

$$K_{\epsilon,p}(r) = (\epsilon+r)^{\frac{p(x)}{2}} + \epsilon^2 r \;\text{ and }\; k_{\epsilon,q}(r) = (\epsilon+r)^{\frac{q(x)}{2}},\ \forall\, r \geq 0. \tag{10}$$

Minimizing the new energy $\mathcal{J}_\epsilon(\cdot)$ in (9) gives rise the following optimality conditions

$$\begin{cases} \Delta\left(k_{\epsilon,p}(|\Delta u|^2)\Delta u\right) - \nabla.(k_{\epsilon,q}(|\nabla u|^2)\nabla u) + \lambda_0 u = \lambda_0 f, & \text{in } \Omega,\\ u_\epsilon = \frac{\partial u_\epsilon}{\partial n} = 0, & \text{on } \partial\Omega, \end{cases} \tag{11}$$

where $k_{\epsilon,p}(r) = K'_{\epsilon,p}(r)$ and $k_{\epsilon,q}(r) = K'_{\epsilon,q}(r)$. The weak formulation of (11) can be defined as follows: Find u_ϵ in $\tilde{H}_0^2(\Omega)$ such that:

$$A_1(u;u,v) + A_2(w;u,v) + \lambda_0\int_\Omega uv\,dx = \ell(v), \quad \forall u,v \in \tilde{H}_0^2(\Omega),$$

where

(12)
$$A_1(w;u,v) = \int_\Omega k_{\epsilon,p}(|\Delta w|^2)\Delta u\Delta v\,dx \;\text{ and }\; A_2(w;u,v) = \int_\Omega k_{\epsilon,q}(|\nabla w|^2)\nabla u\nabla v\,dx.$$

and

$$\ell(v) = \lambda_0\int_\Omega fv\,dx. \tag{13}$$

Remark 1. i) Problem (11) is a reasonable numerical approximation of the non-linear equation (6). It can be seen as a relaxation of problem (6) in the space $\tilde{H}_0^2(\Omega)$ in the same framework of relaxation methods for the total variation energy in $H^1(\Omega)$ (see [13]). In addition, this ensures the $\tilde{H}_0^2(\Omega)$- uniform ellipticity which will guarantee the convergence of the linearization iteration that will be discussed in the next section

ii) The parameter ϵ have also an impact on the computed numerical solution. In fact, it is related to the edges width and should be small enough in order to get sharp edges, see [2, 15] for more details. Moreover, since we have added the term $\epsilon^2\int_\Omega |\Delta u|^2dx$ to energy, the solution belongs to the space $H^2(\Omega)$ and the bi-Laplace diffusion term $\epsilon^2\Delta^2 u$ will appear in the PDEs (11). Thus, this term clearly have a smoothness effect of the solution.

3.1.1. Convergence analysis. In this part, we give using Γ-convergence tools [9], the explanation why we can study the regularization energy (9) directly.

Theorem 3.1. [9] *Let $\mathcal{M}$ be a topological space, endowed with a τ-topology, and let $F_\epsilon, F : X \longrightarrow \mathbb{R}$. If $(F_\epsilon)_{\epsilon>0}$ is a decreasing sequence converging to F pointwise, then F_ϵ Γ-converges to the lower semicontinuous envelope of F in $\mathcal{M}$.*

Let us define on the space X the following energy

$$\mathcal{J}_\epsilon^*(u) = \begin{cases} \mathcal{J}_\epsilon(u), & u \in \tilde{H}_0^2(\Omega), \\ +\infty & \text{Otherwise.} \end{cases} \tag{14}$$

Proposition 2. *Energy $\mathcal{J}_\epsilon^*(\cdot)$ admits a unique minimizer u_ϵ in $\tilde{H}_0^2(\Omega)$. Moreover, the sequence $(u_\epsilon)_{\epsilon>0}$ of solutions of the regularized problems (11) converges, in $\tilde{H}_0^2(\Omega)$, to the solution u of the problem (6) as ϵ goes to 0.*

Proof. Using classical techniques of calculus of variation, we can prove that the energy $\mathcal{J}_\epsilon^*(\cdot)$ admits a unique minimizer u_ϵ in $\tilde{H}_0^2(\Omega)$. Moreover, this minimizer u_ϵ has PDEs (11) as Euler Lagrange optimality conditions Furthermore, let us define the energy

$$\mathcal{J}^*(u) = \begin{cases} \mathcal{J}(u), & u \in \tilde{H}_0^2(\Omega), \\ +\infty & \text{Otherwise.} \end{cases} \tag{15}$$

Then, the energy $(\mathcal{J}_\epsilon^*)_{\epsilon>0}(\cdot)$ decreases and pointwise converges to $\mathcal{J}^*(\cdot)$ when ϵ goes to 0. Thus, $\mathcal{J}_\epsilon^*(\cdot)$ Γ-converges to the lower semi-continuous envelope of $\mathcal{J}_*(\cdot)$ for the X-topology, which is $\mathcal{J}(\cdot)$. Therefore, taking the limit as ϵ goes to 0 and using the properties of the Γ-convergence, we get that the sequence minimizers $(u_\epsilon)_{\epsilon>0}$ of the energy $(\mathcal{J}_\epsilon^*(\cdot))_{\epsilon>0}$, that is also solution of (11) converges to the unique minimizer u of the energy $\mathcal{J}(\cdot)$, which is also solution of (6). □

The relaxed problem (11) is nonlinear and it requires the use of a linearization technique for the numerical computation. The Γ-convergence relaxation ensures the $\tilde{H}_0^2(\Omega)$- uniform ellipticity which will guarantee the convergence of the fixed-point method that we describe now.

4. Fixed-point iterations

To approximate a solution of the nonlinear problem (11), we propose a fixed-point linearization method and we prove the convergence of the sequence of solutions of the linearized problems to the solution of the nonlinear problem (11). We refer the reader interested in such fixed point method to [29, 34].

To approximate a solution of the nonlinear problem (11), we use a fixed-point linearization based method which is summarized as follows:

- Choose an initial guess u_0.
- For $n > 1$, find u_n that solves the linear problem:

$$A_1(u_{n-1}; u_n, v) + A_2(u_{n-1}; u_n, v) + \int_\Omega \lambda_0 u_n v \, dx = \ell(v), \quad \forall v \in \tilde{H}_0^2(\Omega), \tag{16}$$

  until a stopping criterion is satisfied.

where $A_i(\cdot;\cdot,\cdot)$ $(i = 1, 2)$ and $\ell(\cdot)$ are given in (12) and (13), respectively. In order to prove the convergence of the above iterative procedure, we need the following lemma:

Lemma 4.1. *Let $J_1(\cdot)$, $J_2(\cdot)$ be the functional on $\tilde{H}_0^2(\Omega)$ defined by:*

$$J_1(v) = \tfrac{1}{2}\int_\Omega \int_0^{|\Delta v|^2} k_{\epsilon,p}(r)\,dr dx, \quad J_2(v) = \tfrac{1}{2}\int_\Omega \int_0^{|\nabla v|^2} k_{\epsilon,q}(r)\,dr dx.$$

Then, for all u_{n-1} and u_n in X, we have:

$$\tfrac{1}{2}\sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_n) - \tfrac{1}{2}\sum_{i=1}^{2} A_i(u_{n-1}; u_{n-1}, u_{n-1}) \geq \sum_{i=1}^{2} J_i(u_{n-1}) - \sum_{i=1}^{2} J_i(u_n).$$

Proof. The function $Q_p(t) = \int_0^t k_{\epsilon,p}(r)dr$ is concave. In fact, we have:

$$Q_p^{''}(t) = k'_\epsilon(t) \leq 0, \quad \forall t > 0.$$

From the propriety of concave differentiable functions, we have:

$$Q'_p(t_1).(t_2 - t_1) - Q_p(t_2) + Q_p(t_1) \geq 0, \quad \forall t_1, t_2.$$

Similarly, if we define $Q_q(t) = \int_0^t k_{\epsilon,q}(r)dr$ we have:

$$Q'_q(t_1).(t_2 - t_1) - Q_q(t_2) + Q_q(t_1) \geq 0, \quad \forall t_1, t_2.$$

Taking the above two inequalities for $t_1 = |\Delta u_{n-1}|^2$ and $t_2 = |\Delta u_n|^2$, and respectively for $t_1 = |\nabla u_{n-1}|^2$ and $t_2 = |\nabla u_n|^2$ and then taking the sum, we get the proof. $\square$

The convergence of the above iterative method is summarized in the following proposition.

Proposition 3. *Let $(u_n)_{n\geq 0}$ be the sequence of solutions of the linearized problems (16), then the sequence of solutions $(u_n)_{n\geq 0}$ of the linearized problems (16) converges to the unique solution u of the nonlinear problem (11) in $\tilde{H}_0^2(\Omega)$ as n goes to $+\infty$.*

Proof. Let $a(w; u, v) = A_1(w; u, v) + A_2(w; u, v) + \lambda_0 \int_\Omega uv\,dx$ and let u_n be the solution of the linearized problem (16). From the coercivity of the bilinear form $a(u_{n-1}; \cdot, \cdot)$ in $\tilde{H}_0^2(\Omega)$, there exists $\alpha_0 > 0$ such that:

$$\begin{aligned}
&\alpha_0 \|u_n - u_{n-1}\|^2_{\tilde{H}_0^2} \leq a(u_{n-1}; u_n - u_{n-1}, u_n - u_{n-1}) \\
&= \sum_{i=1}^{2} A_i(u_{n-1}; u_n - u_{n-1}, u_n - u_{n-1}) + \lambda_0 \int_\Omega u_n^2\,dx \\
&\quad + \lambda_0 \int_\Omega u_{n-1}^2\,dx - 2\lambda_0 \int_\Omega u_n u_{n-1}\,dx \\
&= \sum_{i=1}^{2} A_i(u_{n-1}; u_{n-1}, u_{n-1}) - 2\sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_{n-1}) + \sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_n) + \\
&\quad + \lambda_0 \int_\Omega u_n^2\,dx + \lambda_0 \int_\Omega u_{n-1}^2\,dx - 2\lambda_0 \int_\Omega u_n u_{n-1}\,dx \\
&= \sum_{i=1}^{2} A_i(u_{n-1}; u_{n-1}, u_{n-1}) - 2\sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_{n-1}) + 2\sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_n) \\
&\quad - \sum_{i=1}^{2} A_i(u_{n-1}; u_n, u_n) + \lambda_0 \int_\Omega u_n^2\,dx + \lambda_0 \int_\Omega u_{n-1}^2\,dx - 2\lambda_0 \int_\Omega u_n u_{n-1}\,dx.
\end{aligned}$$

From the weak formulation (16), we get:

$$2\sum_{i=1}^{2} A_i(u_{n-1};u_n,u_{n-1}) = 2\ell(u_{n-1}) - 2\lambda_0\int_\Omega u_n u_{n-1}\,dx,$$
$$2\sum_{i=1}^{2} A_i(u_{n-1};u_n,u_n) = 2\ell(u_n) - 2\lambda_0\int_\Omega u_n^2\,dx.$$

Therefore, we have:

$$\begin{aligned}\alpha_0\|w_n - u_{n-1}\|^2 \leq & \left(\sum_{i=1}^{2} A_i(u_{n-1};u_{n-1},u_{n-1}) + \lambda_0\int_\Omega u_{n-1}^2\,dx - 2\ell(u_{n-1})\right)\\ & - \left(\sum_{i=1}^{2} A_i(u_{n-1};u_n,u_n) + \lambda_0\int_\Omega u_n^2\,dx - 2\ell(u_n)\right)\end{aligned}$$

The above inequality, together with Lemma 4.1 give:

$$\begin{aligned}0 < \alpha_0\|u_n - u_{n-1}\|^2 \leq & \left(\sum_{i=1}^{2} J_i(u_{n-1}) + \lambda_0\int_\Omega u_{n-1}^2\,dx - 2\ell(u_{n-1})\right)\\ & - \left(\sum_{i=1}^{2} J_i(u_n) + \lambda_0\int_\Omega u_n^2\,dx - 2\ell(u_n)\right)\\ =& 2(\mathcal{J}_\epsilon(u_{n-1}) - \mathcal{J}_\epsilon(u_n)).\end{aligned}$$

Therefore, we have:

$$0 < \alpha_0\|u_n - u_{n-1}\|^2 \leq 2(\mathcal{J}_\epsilon(u_{n-1}) - \mathcal{J}_\epsilon(u_n)), \tag{17}$$

where the energy $\mathcal{J}_\epsilon(\cdot)$ is given in (9). Inequality (17) implies that the sequence $(\mathcal{J}_\epsilon(u_n))_n$ is monotone decreasing. Moreover, the energy $\mathcal{J}_\epsilon(\cdot)$ has unique minimizer, which is solution of (11), which means that the sequence $(\mathcal{J}_\epsilon(u_n))_n$ is bounded below, and then that it converges. From inequality (17), we deduce that $\|u_n - u_{n-1}\|^2_{\tilde{H}_0^2}$ converges to 0 as n goes to $+\infty$.

Now, let u be the unique weak solution of the nonlinear problem (16). We consider the nonlinear operator

$$B(u) = \Delta(k_\epsilon(|\Delta u|^2)\Delta u) + \nabla\cdot(k_\epsilon(|\nabla u|^2)\nabla u) + \lambda_0 u.$$

Then, from the coercivity of $B(u)$ in $\tilde{H}^2_0(\Omega)$, we get:

$$\begin{aligned}
\alpha_0\|u_{n-1}-u\|^2_{\tilde{H}^2_0} &\leq \langle Bu_{n-1}-Bu, u_{n-1}-u\rangle_2 \\
&= \sum_{i=1}^{2} A_i(u_{n-1};u_{n-1},u_{n-1}-u) - \sum_{i=1}^{2} A_i(u;u,u_{n-1}-u) \\
&= \sum_{i=1}^{2} A_i(u_{n-1};u_{n-1}-u_n,u_{n-1}-u) - \sum_{i=1}^{2} A_i(u;u,u_{n-1}-u) \\
&+ \sum_{i=1}^{2} A_i(u_{n-1};u_n,u_{n-1}-u) \\
&= \sum_{i=1}^{2} A_i(u_{n-1};u_{n-1}-u_n,u_{n-1}-u) - \ell(u_{n-1}-u) + \ell(u_{n-1}-u) \\
&= \sum_{i=1}^{2} A_i(u_{n-1};u_{n-1}-u_n,u_{n-1}-u).
\end{aligned}$$

Hence, from the continuity of $A_i(u_{n-1};\cdot,\cdot)$ we write:

$$\|u_{n-1}-u\|^2_{\tilde{H}^2_0} \leq c\|u_{n-1}-u_n\|_{\tilde{H}^2_0}\|u_{n-1}-u\|_{\tilde{H}^2_0}, \;\; c>0,$$

which implies that $(u_n)_n$ converges to u which is the unique weak solution of the nonlinear problem (16). $\square$

Remark 2. We note that in practise, the homogeneous Dirichlet boundary condition, $u = 0$ on $\partial\Omega$ is not convenient and not always satisfied. As the image f is noisy, we consider the condition $u = f^\star$ on $\partial\Omega$ where $f^\star$ is a smoothed version of f. The theoretical analysis is still applicable by using lifting techniques using a lift $u_{f^\star}$ such that $u_{f^\star} = f^\star$ and $\frac{\partial u_{f^\star}}{\partial n} = 0$ on $\partial\Omega$.

5. Morley Finite-element discretization and adaptivity

Let $\mathbb{P}_2$ be the space of all polynomials with degree not greater than 2. We assume that the domain Ω is polygonal and we consider a regular family of triangulations $\mathcal{T}^h$ made of element which are triangles with a maximum size h, satisfying the usual admissibility assumptions, i.e. the intersection of two different elements is either empty, a vertex, or a whole edge. Then, given a triangle K, its three vertices is denoted by a_j, $1 \leq j \leq 3$. The edge of K opposite to a_j is denoted by F_j, $1 \leq j \leq 3$. Denote the measures of K and F_j by $|K|$ and $|F_j|$, respectively. Morley element can be described by (K, P_K, Φ_K) with

1. K is a triangle.
2. $P_K \in \mathbb{P}_2(K)$ where $\mathbb{P}_2(K)$ is the space of all polynomials with degree not greater than 2 on K.
3. Φ_K is the vector of degrees of freedom whose components are:

$$v(a_j), \;\; \frac{1}{|F_j|}\int_{F_j}\frac{\partial v}{\partial \nu}ds, \; 1 \leq j \leq 3, \;\; \forall v \in \mathbb{P}_2.$$

Let $\mathcal{V}_{int}$ (resp. $\mathcal{V}_{ext}$) be the set of internal vertices (resp. vertices on the boundary) of $\mathcal{T}^h$. We also denote by $\mathcal{E}(\Omega)$ (resp. $\mathcal{E}(\partial\Omega)$) set of all internal edges (resp. boundary edges) of $\mathcal{T}^h$. Then, we define the finite element space corresponding to the Morley element defined above as follows

$$X^h = \Big\{ v|_K \in \mathbb{P}_2 \text{ such that } v \text{ is continuous at } \mathcal{V}_{int} \text{ and vanishes at } \mathcal{V}_{ext}, \tag{18}$$
$$\forall F \in \mathcal{E}(\Omega), \int_F \left[\frac{\partial v}{\partial n}\right]_F ds = 0; \text{ and } \forall F \in \mathcal{E}(\partial\Omega), \int_F \frac{\partial v}{\partial n} ds = 0 \Big\},$$

$[\phi]_F$ is the jump of ϕ across the interior edge F.

Given an initial function $u_0^h \in X^h$ the discrete problem is written as follows: For $n > 0$, find $u_n^h \in X^h$, such that:

$$A_1(u_{n-1}^h; u_n^h, v^h) + A_2(u_{n-1}^h; u_n^h, v^h) + \int_\Omega u_n^h v^h \, dx = \ell(v^v), \quad \forall v^h \in X^h, \tag{19}$$

where $A_1(\cdot;\cdot,\cdot)$ and $A_2(\cdot;\cdot,\cdot)$ are defined in (12) and $\ell(\cdot)$ is defined in (13).

5.1. Adaptive strategy. For the discretization, we will use a discontinuous approximation of $p(\cdot)$ and $q(\cdot)$. Therefore, we suppose now that the domain Ω is partitioned into I disjoint sub-domains $(\Omega_i)_i$ such that p and q are given by the piecewise constant scalar function:

$$p = p_i \text{ and } q = q_i, \text{ in } \Omega_i, \quad K = 1, \dots, I,$$

where $p_m = \min_{1 \le K \le I} p_K > p^- = 1$ and $p_M = \max_{1 \le K \le I} p_K \le p^+ = 2$ (q_m and q_M are defined similarly).

Then, we consider a multiscale spatially adaptive choice for variable exponents $p(\cdot)$ and $q(\cdot)$ with the help of the maps furnished by a local error indicators. The approach that we propose automatically balances between L^1 and L^2-regularization effects in the spirit of the works proposed in [8, 16, 42, 23]. This choice is in accordance with anisotropic PDE-based methods which are largely used in order to overcome over smoothing and staircasing artefacts. It is based on a posteriori indicator map which is constructed using information that comes from a smoothed structure tensor (SST). The latter has been extensively used in nonlinear anisotropic PDEs [47, 49].

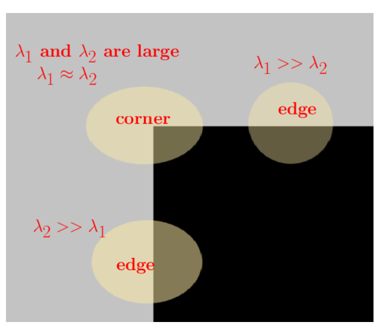


FIGURE 1. The smoothed structure tensor (SST) detector summary. The error indicator that we use in this work favours better the effect of anisotropy. In fact, in regions where $\lambda_+^1 \approx \lambda_-^1 \approx 0$, there are very few variation and the region does not contain any edges or corners, i.e., almost flat. In regions where $\lambda_+^1 >> \lambda_-^1 \approx 0$, there are a lot of variations and the current position represents edges. Elsewhere, we have $0 \ll \lambda_-^1 \le \lambda_+^1$, and we are located on a corner structure in the image (see [47, 49]).

5.1.1. Multiscale choice. Let u^h be the discrete solution (image), we define three tensor $K^1_\sigma(u^h)$, $K^y_\sigma(u^h_x)$ and $K^y_\sigma(u^h_x)$, where for a scalar function u, we have

$$K_\sigma(u) = G_\sigma \star (\nabla u(\nabla u)^T) = \begin{bmatrix} S_\sigma \star (u_x)^2 & S_\sigma \star u_x u_y \\ S_\sigma \star u_y u_x & S_\sigma \star (u_y)^2 \end{bmatrix}.$$

The function $S_\sigma(x) = (\sigma\sqrt{2\pi})^{-1}\exp(-|x|^2/\sigma^2)$ is a 2D Gaussian kernel and $\star$ denotes the convolution operator which makes the structure tensor exceedingly robust to noise. This structure tensor was introduced in [47, 49] and is considered so far very sufficient to identify discontinuities (edges). It has been extensively used in nonlinear anisotropic PDEs. Most importantly, the matrix $K_\sigma(u)$ is positive semi-definite and has two eigenvalues $(\lambda^1_+, \lambda^1_-)$ which are well adapted to discriminate different geometric features at different scales i.e. edge, corner, thin structure and flat regions of the image, see Fig. 1.

To perform the adaptive local choice of $p(\cdot)$ and $q(\cdot)$, we use the tensors $K^1_\sigma(u^h)$, $K^y_\sigma(u^h_x)$ and $K^y_\sigma(u^h_x)$ and their eigenvalues $(\lambda^1_+, \lambda^1_-)$, $(\lambda^2_+, \lambda^2_-)$ and $(\lambda^3_+, \lambda^3_-)$, respectively. We start by the selection of the exponent q which is linked the term ∇u^h. For its selection, we use the eigenvalues $(\lambda^1_+, \lambda^1_-)$ of the tensor $K^1_\sigma(u^h)$ as it is able to detect the variation of ∇u^h. However, it is unable do detect the discontinuities of the first derivatives u^h_x and u^h_y, i.e. regions where u^h continuous but its first derivatives u^h_x and u^h_y aren't. Such regions are certainly of interest in denoising as they encode some image structures and features. Thus, for the selection of p, which is the exponent of the high-order term, we use the eigenvalues $(\lambda^2_+, \lambda^2_-)$ and $(\lambda^3_+, \lambda^3_-)$ of the tensor $K^y_\sigma(u^h_x)$ and $K^y_\sigma(u^h_x)$ as they can detect discontinuities of u^h_x and u^h_y.

The use of the eigenvalues $(\lambda^i_+, \lambda^i_-)$ $(i = 1, 2, 3)$ allows to automatically self-tune the exponents of the regularization terms according to local image structures. Furthermore, the new model will have anisotropic nonlinear diffusion property that comes from these eigenvalues which encode some anisotropy properties. Moreover, the model acts in a multiscale level by adopting the diffusion dependant on the region. It will create a diffusion process in three different levels that depend on the different scales in the image, i. e. homogeneous, edges and discontinuities. We also use a mesh refinement in order to allow the model to capture most of the image features even those of small scales, e. g. see [3, 4, 45]

Consequently, the proposed model will have multiscale character that will help to effectively reduce noise and keep the important feature sharp.

The overall adaptivity steps are summarized in the following algorithm:

Adaptivity steps

(1) Consider $p_0 = q_0 = 2$ and compute u^h_0 solution of (19) in the initial grid $\mathcal{T}^h_0$.
(2) **Adaptive steps: For $k > 0$:**
   - Compute u^h_k on $\mathcal{T}^h_k$ with $p = p_k$ and $q = q_k$ using fixed-point iterations.
   - Build an adapted mesh $\mathcal{T}^h_{k+1}$ (in the sense of the finite element method, i.e., with respect to the parameter h) with a metric error indicator.
   - Perform a local choice of $p(x)$ and $q(x)$ on $\mathcal{T}^h_{k+1}$ to obtain new functions p_{k+1} and q_{k+1}.
(3) If $||u^h_{k+1} - u^h_k|| \leq tolerance$ or $n \geq n_{max}$, stop. Otherwise go to Step 2.

During the adaptation, we use the following formula for each triangle K of the mesh:

$$q^K_{k+1} = \max\left(\frac{q^K_k}{1+\tau*\left(\left(\frac{\lambda^1_-}{\|\lambda^1_-\|_\infty}\right)-\mu\right)^+ +\tau*\left(\left(\frac{\lambda^1_+}{\|\lambda^1_+\|_\infty}\right)-\mu\right)^+},1\right), \tag{20}$$

where τ is a coefficient chosen to control the rate of decrease in $q(\cdot)$ and $u^+ = \max(u,0)$. For the choice of $p(\cdot)$, both $K^x_\sigma(u^h_x)$ and $K^y_\sigma(u^h_x)$ quantities will be involved as they contain information about the image derivatives. Thus, we assume that we can write $p = \frac{p^2+p^3}{2}$ where p^2 and p^2 will be computed separately from $(\lambda^2_+,\lambda^2_-)$ and $(\lambda^3_+,\lambda^3_-)$, respectively. We start by $p^2_0 = p^3_0 = 2$ and we update them following the formula

$$p^i_{k+1} = \max\left(\frac{p^i_k}{1+\tau*\left(\left(\frac{\lambda^i_-}{\|\lambda^i_-\|_\infty}\right)-\mu\right)^+ +\tau*\left(\left(\frac{\lambda^i_+}{\|\lambda^i_+\|_\infty}\right)-\mu\right)^+},1\right), i=2,3. \tag{21}$$

After that, we set the new exponent $p_{k+1} = \frac{p^2_{k+1}+p^3_{k+1}}{2}$. The update formula of $p(\cdot)$ and $q(\cdot)$ are explained as follows:

- *Flat/noisy regions in the image*: The first level of diffusion is in flat regions where $\lambda^1_+ \approx \lambda^1_- \approx 0$. In this case, the denominator in the formula (20) is close to 1 which will keep q^K_{k+1} close to 2. Note also in these regions, $\lambda^i_+ \approx \lambda^i_- \approx 0$ $(i=2,3)$ and p^K_{k+1} will be close to 2. The model behaves like biharmonic equation leading to strong smoothness that effectively damps oscillation.
- *Edge and derivatives discontinuities*: The second level of diffusion is near the edges of the image and its derivatives. In these regions, the eigenvalue λ^1_- is very small and therefore the second quantity in the denominator will be close to 0. However, the variation the second eigenvalue λ^1_+ in these regions is very important as λ^1_+ usually detects both edges and corner. Thus, the quantity $\frac{\lambda^1_+}{\|\lambda^1_+\|_\infty}$ will be very important in the selection of q^K_{k+1}. The analysis is the same for the selection of p^K_{k+1} in the formula (21), i.e. λ^i_- is very small and λ^i_+ is important near the near edges/discontinuities of the derivatives u^h_x and u^h_y. In this case, we slow-down diffusion and the rate of decreasing of $q(\cdot)$ and $p^i(\cdot)$ in the formulae (20) and (21) is the same rate of:

$$\frac{q_k}{1+\tau*\left(\left(\frac{\lambda^1_+}{\|\lambda^1_+\|_\infty}\right)-\mu\right)^+} \quad \text{and} \quad \frac{p^i_k}{1+\tau*\left(\left(\frac{\lambda^i_+}{\|\lambda^i_+\|_\infty}\right)-\mu\right)^+},\ i=2,3.$$

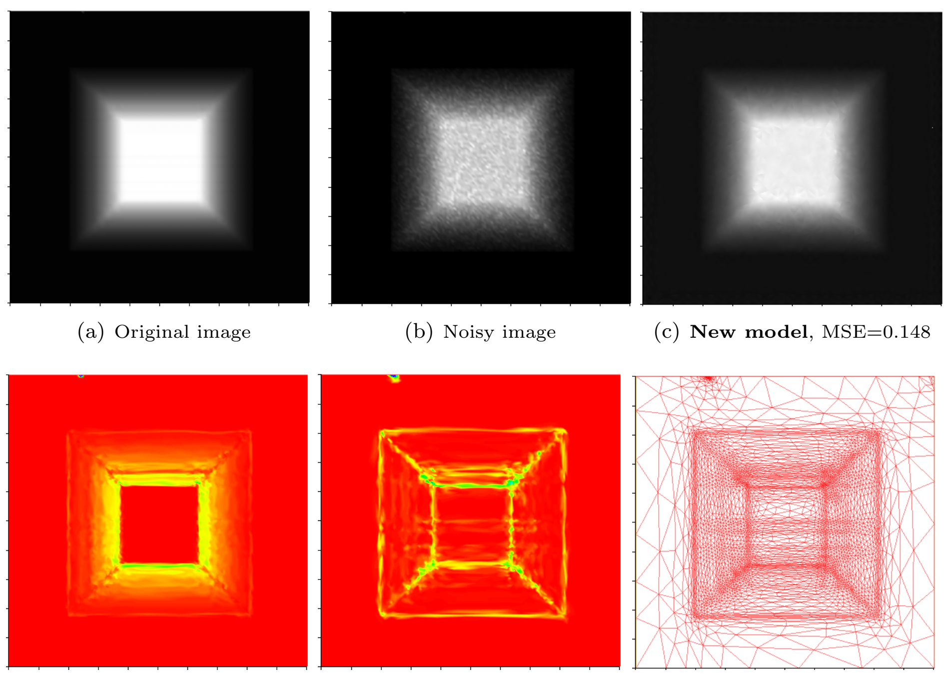
- *Corners in the image or its derivatives*: Near a corner, $0 \ll \lambda^i_- \leq \lambda^i_+$ $(i=1,2,3)$ and the two quantities in the denominators in (20) and (21) are both high. Therefore, the exponents q_{k+1} and p_{k+1} will be close to 1 and corner points are not smoothed out. In addition, the exponents q_{k+1}

and p_{k+1} will decease near the corners faster than near edges which gives rise to a third level of diffusion.

To conclude this section, we notice that it may appear "exagerated" to change $p(\cdot)$ and $q(\cdot)$, hence the model, at the scale of the elements of the mesh. However, in practice the tight location of the singularities yields a thin zone surrounding the singular set where $p(\cdot)$ and $q(\cdot)$ are close to 1 and the rest where $p(\cdot) = q(\cdot) = 2$.

6. Numerical experiments

In this work, all the PDEs are solved with the open source software FreeFem++ [30]. We present several numerical examples, in image restoration problems for additive noise removal. In all examples, we set $\lambda_0 = 10^5$, $\epsilon = 10^{-05}$ and $\mu = 10^{-1}$. For the fixed point algorithm, we choose an initial condition $u_0 = 1$ for all the examples.



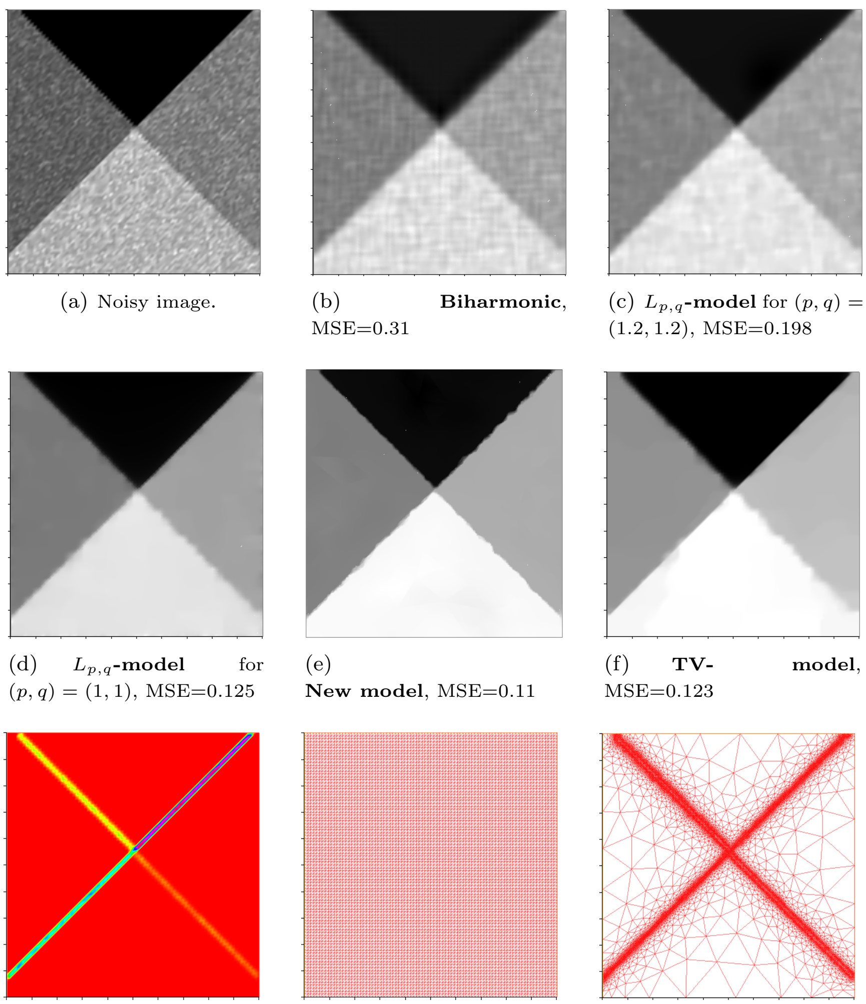
(a) Original image (b) Noisy image (c) **New model**, MSE=0.148

(d) The map of $\lambda_1^1 + \lambda_2^1$ (e) The map of $\sum_{i=2}^{3} \lambda_1^i + \lambda_2^i$ (f) Final mesh

FIGURE 2. For the map $\lambda_1^1 + \lambda_2^1$ in Fig 2-(d), it is clear that the tensor $K_\sigma^1(u^h)$ is unable to detect discontinuities of second kind, i.e., discontinuities of the gradient. However, the tensors $K_\sigma^2(u_x^h)$ and $K_\sigma^2(u_y^h)$ give a good description of this discontinuities as they include some informations about high-order derivatives. This can bee easy seen from the map of $\sum_{i=2}^3 \lambda_1^i + \lambda_2^i$ in Fig 2-(e).

For comparison, we include 4 methods on the comparison ist. We denote by "**New model**" our adaptive model, "**Biharmonic model**" the model (11) for $p = q = 2$, "$L_{p,q}$**-model**" the model (11) for a constant $p, q \in [1, 2)$ and "**TV-model**" the total variation model. All the models are qualitatively compared by giving the mean square error values (MSE) or/and the Peak signal-to-noise ratio ($PSNR$), where a higher $PNSR$ value implies a better denoising quality. For the adaptation, we recall that we start the computations with exponents $p(x) = q(x) \equiv 2$ in all the image domain Ω. During the adaptation steps, the exponents decrease in high gradient zones (formally close to 1) encouraging possible jumps in these areas.

Example 1. In Fig. 2, we test the adaptive model for a synthetic image of resolution 128×128 which contains discontinuities of first- and second-order, i.e., of u and ∇u. We display the denoised image and the different eigenvalues λ_1^i and λ_2^i $(i = 1, 2, 3)$. Because it contains only information about first-order derivatives, λ_1^1 and λ_2^1 are unable to detect second-order geometrical features such as points of gradient discontinuity. Such points correspond to the crease discontinuities, i.e. the points where u is continuous while ∇u is discontinuous. However, the output image for λ_1^i and λ_2^i $(i = 2, 3)$ are given by meaningful boundaries which correspond to discontinuities set of the image and of its first derivatives. This example clearly shows the benefits of using a tensor for the derivatives, i.e, $K_\sigma^x(u_x^h)$ and $K_\sigma^y(u_y^h)$.



(a) Noisy image.

(b) **Biharmonic**, MSE=0.31

(c) $L_{p,q}$**-model** for $(p, q) = (1.2, 1.2)$, MSE=0.198

(d) $L_{p,q}$**-model** for $(p, q) = (1, 1)$, MSE=0.125

(e) **New model**, MSE=0.11

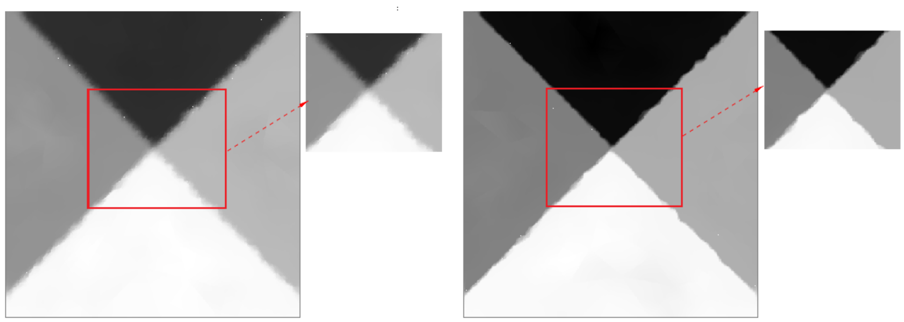
(f) **TV-model**, MSE=0.123

(g) The map of $p(\cdot) + p(\cdot)$

(h) Initial mesh

(i) Final mesh

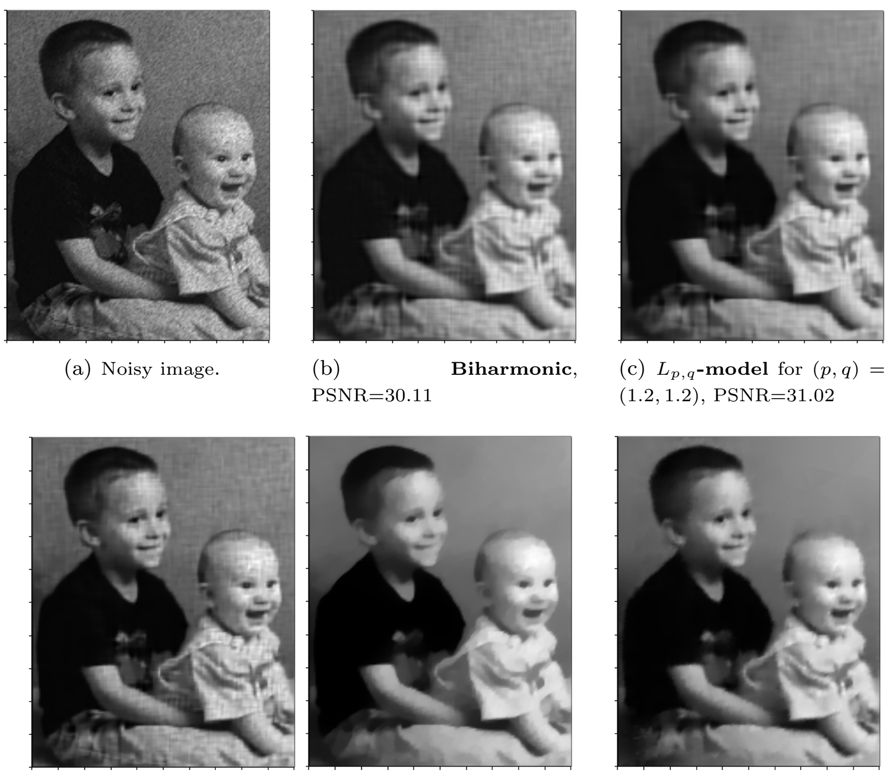
FIGURE 3. Comparison of the different models.



(a) Denoised image for $\epsilon = 0.05$.

(b) Denoised image for $\epsilon = 0.0005$.

FIGURE 4. Denoised image using **New model** for $\epsilon = 0.05$ and $\epsilon = 0.0005$. It is clear from the zoomed regions that smaller ϵ leads to shaper edges.



(a) Noisy image.

(b) **Biharmonic**, PSNR=30.11

(c) $L_{p,q}$-**model** for $(p, q) = (1.2, 1.2)$, PSNR=31.02

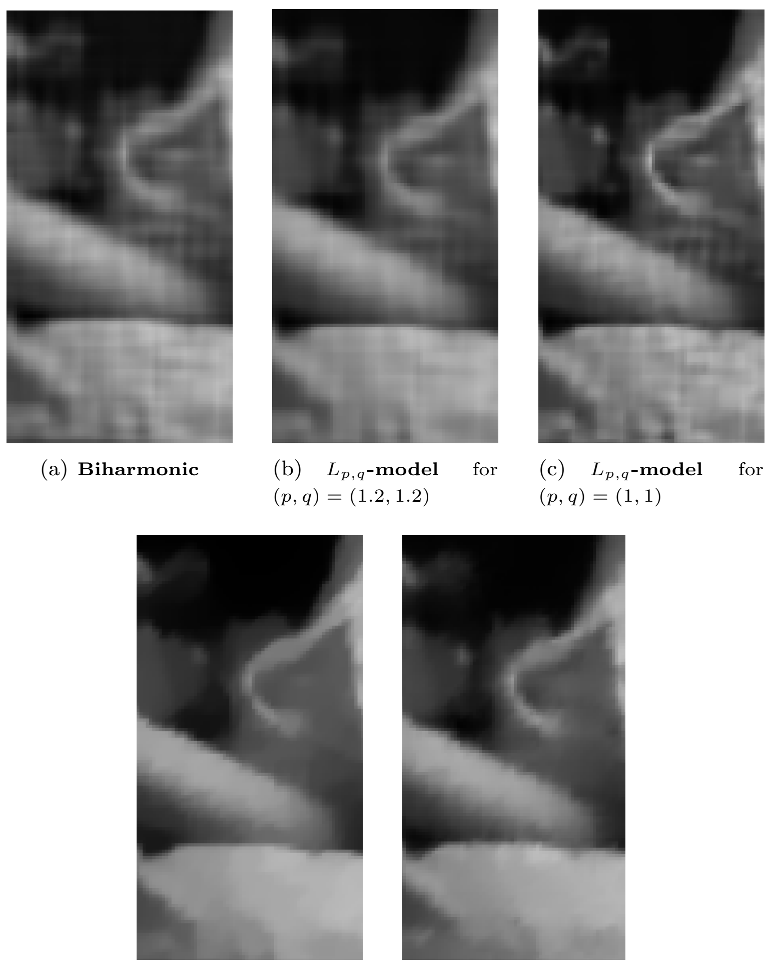
(d) $L_{p,q}$-**model** for $(p, q) = (1, 1)$, PSNR=31.42

(e) **New model**, PSNR=32.53

(f) **TV-model**, PSNR=32.24

FIGURE 5. Comparison of the different models for real image denoising. The initial image is corrupted with a Gaussian noise having zero mean and standard deviation 15.

Example 2. In Fig. 3, we display the denoised image of resolution 128×128 using the different models, i.e., our model, the "**Biharmonic model**", "$L_{p,q}$-**model**" for $(p, q) \in \{(1, 1), (1.2, 1.2)\}$ and the "**TV-model**". It is clear that the result obtained by the proposed method is visually better than those of the "**Biharmonic model**", "$L_{p,q}$**-model**" which produce a blurred edges. However, the proposed model and the $L_{p,q}$**-model**" produce sharp edges and clean image. We also display the map furnished by the $p(\cdot) + q(\cdot)$ in Fig. 1(e) and which shows that it is essentially equivalent to edge detection function. For this example, we also investigate the choice of the parameter ϵ by testing two different values, i.e. $\epsilon = 0.05$ and $\epsilon = 0.005$. It is clear from the zoom captions in Fig. 4 that smaller ϵ leads to sharper edges.



(a) **Biharmonic**

(b) $L_{p,q}$**-model** for $(p, q) = (1.2, 1.2)$

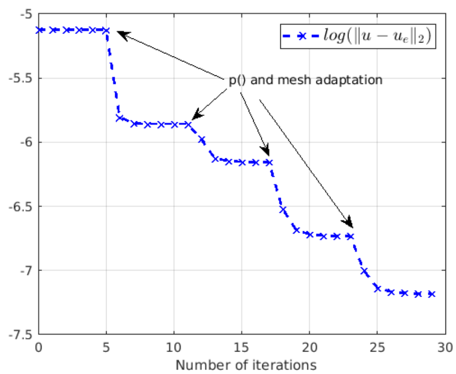
(c) $L_{p,q}$**-model** for $(p, q) = (1, 1)$

(d) **New model**

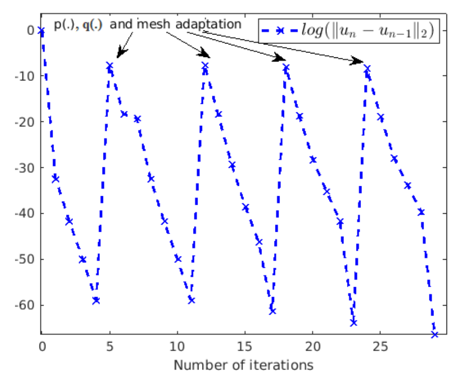
(e) **TV-model**

FIGURE 6. Zoom captions from the test in Fig. 5.

Example 3. In Fig. 5, we demonstrate the effectiveness of the proposed model applied to real noisy image of resolution 200×300 corrupted with Gaussian noise having zero mean and standard deviation $std = 15$. We compare all the competitive models by giving the $PSNR$ values. We display the restored images which show that the proposed **New model** performs better than "**Biharmonic model**", "$L_{p,q}$**-model**" for $(p, q) \in \{(1, 1), (1.2, 1.2)\}$ and the "**TV-model**". The difference



(a)



(b)

FIGURE 7. (a): L^2-error $e_r = \|u - u_e\|_2$ between the denoised image u and the original clean image u_e in semi-log scale. (b) the curve of the relative error $\|u^n - u^{n-1}\|_2/\|u^{n-1}\|_2$ in semi-log scale.

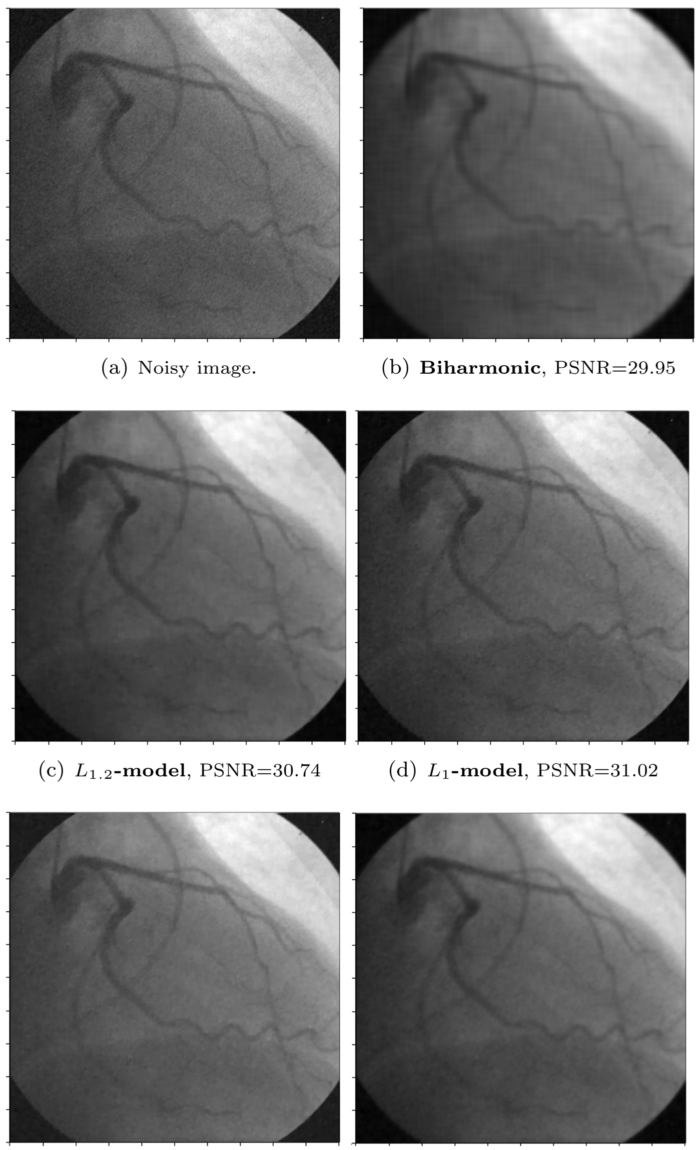
between all results can be seen by zooming some parts of the images in Fig. 6 where we easily notice the advantage of the **New model**. In fact, the latter gives a satisfactory restoration and it behaves like a linear model in smooth parts and as a nonlinear one near edges. We also display the adapted mesh obtained at the end of the adaptive process. The final mesh is refined and fitted to the geometry of the image.

TABLE 1. Runtime (in seconds) comparison between all models in denoising examples 1-4.

	Examples			
	Example 1	Example 2	Example 3	Example 4
Biharmonic model	2.5	2.4	3.62	4.51
$L_{1.2,1.2}$**-model**	28.12	28.43	39.15	40.12
$L_{1,1}$**-model**	28.45	28.89	39.78	40.26
TV-model	20.14	21.57	24.15	29.47
New model	20.87	21.96	22.17	24.85

For this example, we performed 5 iterations for $p(\cdot), q(\cdot)$ and the mesh adaptation, where for each adaptation iteration, 6 inner iteration were performed for the fixed point algorithm. We give show in Fig. 7- (a) the curve of L^2-error $e_r = \|u - u_e\|_2$ between the denoised image u and the original clean image u_e as function of the overall all 5×6 iterations. The curve monotonically decreases which proves the convergence of the algorithm. Moreover, in order to show the convergence of the fixed point linearization algorithm, we show the curve of the relative error $\|u^n - u^{n-1}\|_2/\|u^{n-1}\|_2$ as function of the fixed point iteration in the 5 adaptation iterations. The curve is cyclic where we clearly see 5 cycles that correspond to the fixed point iterations in one adaptation iteration. The curve monotonically decreases for each cycle confirming the convergence of the fixed point algorithm.

Example 4. In Fig. 8, we illustrate the efficiency of the proposed model's ability to restore a medical noisy image of resolution 300×300 which contains thin structures. The image is corrupted with Gaussian noise having zero mean and



(a) Noisy image. (b) **Biharmonic**, PSNR=29.95

(c) $L_{1.2}$**-model**, PSNR=30.74 (d) L_1**-model**, PSNR=31.02

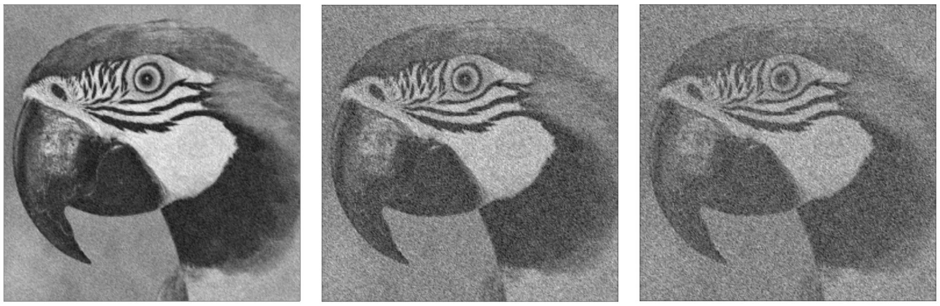
(e) **New model**, PSNR=32.82 (f) **TV-model**, PSNR=32.34

FIGURE 8. Medical image denoising: Comparison between the different models.

different standard deviation $std = 10$. We display the denoised images using the different approaches and we give their correspondent $PSNR$ values.

Different noise levels. In Fig. 9, we demonstrate the effectiveness of the proposed model applied to three noisy images corrupted with Gaussian noise having zero mean and different standard deviations $std = 17, 25$ and 35. Our model seems to be stable w.r.t to noise up to $std = 35$. This is in concordance wit most of PDE approaches which do not give very satisfactory results for very high noise level.

Runtime. We give a runtine comparison for all tested models in Table 1. The **Biharmonic model** is faster than other models as it is linear and does not require



(a) Noisy image, $\sigma = 17$ (b) Noisy image, $\sigma = 25$ (c) Noisy image, $\sigma = 35$

(d) Restored image, $PSNR = 30.02$ (e) **Our model**: Restored image, $PSNR = 28.32$ (f) **Our model**: Restored image, $PSNR = 23.86$

Figure 9. Image restoration for different level of Gaussian noise. First row: Input noisy image obtained by adding Gaussian noise with zero mean and standard deviation 17, 25 and 35, respectively. Second column: Restored images.

Figure 10. The set12 dataset.

any linearization iterations, unlike the other models. Moreover, our **New model** which is of fourth-order, requires almost the same time as the **TV-model** even the

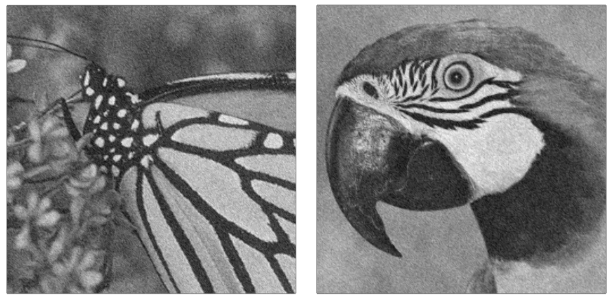


FIGURE 11. The noisy Starfish and Parrot images.

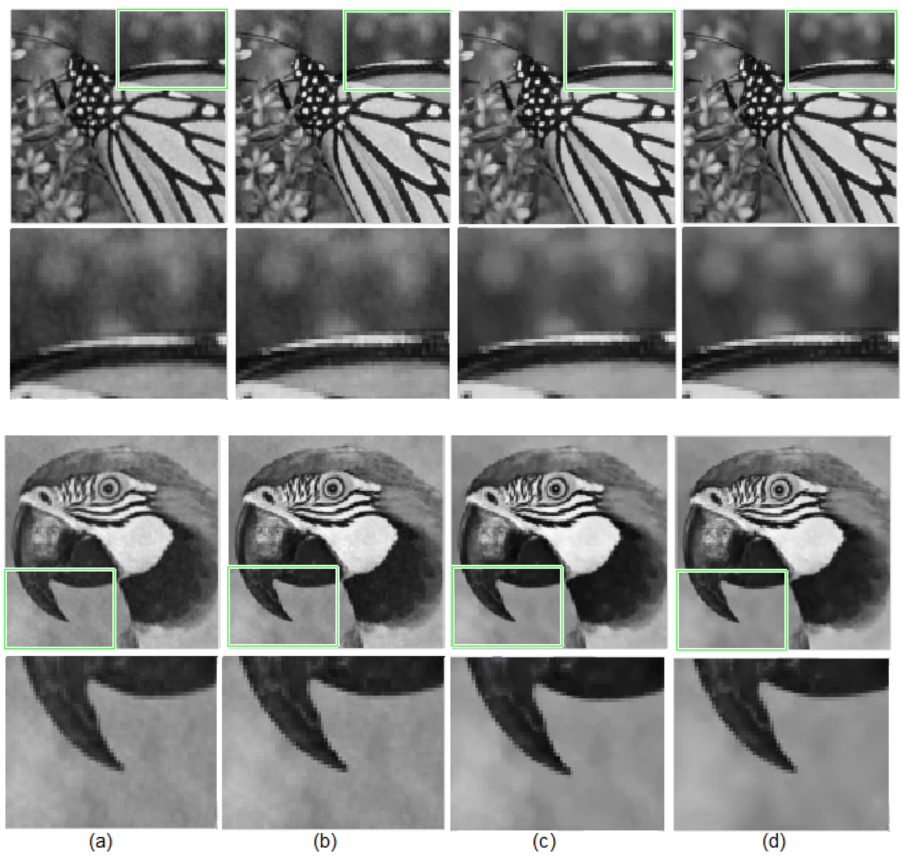


FIGURE 12. Comparison of the different models in denosing Starfish and Parrot images. First and thrid rows, form left to right: Denoised images using **Biharmonic model**, $L_{p,q}$**-model** for $(p, q) = (1.2, 1.2)$, **TV-model** and **New model**. Second and fourth rows: zoom captions from the denosied images. It is visually clear that **New model** gives better result.

latter is a second-order PDE. The runtime of **New model** without mesh adaptation will be the same as $L_{1.2,1.2}$**-model** and $L_{1,1}$**-model**. This clearly emphasises the importance of mesh adaptation in speeding up the runtime.

Set12 dabaset: To test the effectiveness of the proposed approach on 12 images of resolution 256×256 from Set12 dataset, see Fig. 10. All images were corrupted with Gaussian noise having zero mean and standard deviation $std = 15$. We compare with the discussed competitive models (See Fig. 11 and Fig. 12) and summarize the results in Table 2. It can be seen that PSNR results for all methods are comparable, with a small advantage for **New model**.

Table 2. Comparison between all models in denoising the 12 images form Set12 dataset.

	C.man	House	Peppers	Starfish	Monarch	Airplane
Biharmonic model	30.15	29.45	30.63	29.35	29.51	30.68
$L_{1.2,1.2}$**-model**	31.26	30.75	31.53	30.45	30.61	31.78
TV-model	32.10	31.09	31.73	30.55	30.74	32.02
New model	32.68	32.23	32.47	32.15	31.81	32.87
	Parrot	Lena	Barbara	Boat	Man	Couple
Biharmonic model	29.36	30.12	30.33	29.69	30.71	30.04
$L_{1.2,1.2}$**-model**	30.24	30.18	30.67	30.19	30.94	30.51
TV-model	32.09	31.87	32.22	32.14	31.85	32.01
New model	32.04	32.17	32.29	32.27	31.93	32.12

7. Conclusion

Image denoising is a challenging modelling task with a broad range of applications, in particular in medical imaging. The work presented in this paper deals with the image denoising problem. We have considered a multiscale approach for building a nonstandard high-order and anisotropic model, based on the $(p(\cdot), q(\cdot))$-Kirchhoff operator. We discussed a practical and efficient strategy for the choice of the exponents $1 < p(\cdot), q(\cdot) \leq 2$, locally and adaptively, which allows us to solve in respect of the fine spatial scales. We analysed the proposed model and the numerical algorithm employed. Numerical realisations have shown the proposed method out-performs the compared classical approaches.

References

[1] R. Aboulaich, D. Meskine, and A. Souissi, New diffusion models in image processing, Computers & Mathematics with Applications, 56 (2008), pp. 874–882.
[2] R. Acar and C. R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, Inverse problems, 10 (1994), p. 1217.
[3] Z. Belhachmi and F. Hecht, Control of the effects of regularization on variational optic flow computations, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 1–19.
[4] Z. Belhachmi, M. Kallel, M. Moakher, and A. Theljani, Weighted harmonic and complex Ginzburg-Landau equations for gray value image inpainting, International Journal of Numerical Analysis & Modeling, 13 (2016), pp. 782–801.
[5] M. Belloni and B. Kawohl, A direct uniqueness proof for equations involving the p-Laplace operator, Manuscripta Mathematica, 109 (2002), pp. 229–231.
[6] K. Ben Haddouch, Z. El Allali, N. Tsouli, S. El Habib, and F. Kissi, Existence of solutions for a fourth order eigenvalue problem with variable exponent under Neumann boundary conditions, Bol. Soc. Paran. Mat., 34 (2016), pp. 253–272.
[7] M. Benning, C. Brune, M. Burger, and J. Müller, Higher-order TV methods enhancement via Bregman iteration, Journal of Scientific Computing, 54 (2013), pp. 269–310.
[8] P. Blomgren, T. Chan, P. Mulet, and C. Wong, Total variation image restoration: numerical methods and extensions, in Image Processing, 1997. Proceedings., International Conference on, vol. 3, 1997, pp. 384–387.

[9] A. Braides, *Gamma*-Convergence for Beginners, no. 22, in Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2002.

[10] K. Bredies, K. Kunisch, and T. Pock, Total generalized variation, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526.

[11] K. Bredies, K. Kunisch, and T. Pock, Total generalized variation, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526.

[12] A. Chambolle and P. Lions, Image recovery via total variation minimization and related problems, Numerische Mathematik, 76 (1997), pp. 167–188.

[13] A. Chambolle and P.-L. Lions, Image recovery via total variation minimization and related problems, Numerische Mathematik, 76 (1997), pp. 167–188.

[14] T. Chan, A. Marquina, and P. Mulet, High-order total variation-based image restoration, SIAM Journal on Scientific Computing, 22 (2000), pp. 503–516.

[15] Q. Chang and I.-L. Chern, Acceleration methods for total variation-based image denoising, SIAM Journal on Scientific Computing, 25 (2003), pp. 982–994.

[16] Y. Chen, S. Levine, and M. Rao, Variable exponent, linear growth functionals in image restoration, SIAM Journal on Applied Mathematics, 66 (2006), pp. 1383–1406.

[17] D. Cruz-Uribe and A. Fiorenza, Introduction to the variable Lebesgue spaces, in Variable Lebesgue spaces and hyperbolic systems, Adv. Courses Math. CRM Barcelona, Birkhäuser/Springer, Basel, 2014, pp. 1–90.

[18] R. Dautray and J.-L. Lions, Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 2, Functional and Variational Methods, Springer-Verlag, 1988.

[19] L. Diening, P. Harjulehto, P. Hästö, and M. Ruzicka, Lebesgue and Sobolev spaces with variable exponents, Springer, 2011.

[20] D. Edmunds and J. Rákosník, Sobolev embeddings with variable exponent, Studia Mathematica, 143 (2000), pp. 267–293.

[21] X. Fan and D. Zhao, On the spaces $L^{p(x)}$ and $W^{m,p(x)}$, Journal of Mathematical Analysis and Applications, 263 (2001), pp. 424–446.

[22] X. Fan and D. Zhao, On the spaces $L^{p(x)}(\Omega)$ and $W^{m,p(x)}(\Omega)$, Journal of Mathematical Analysis and Applications, 263 (2001), pp. 424–446.

[23] L. Fang, L. Zhibin, and P. Ling, Variable exponent functionals in image restoration, Applied Mathematics and Computation, 216 (2010), pp. 870 – 882.

[24] R. Filippucci, P. Pucci, and F. Robert, On a p-Laplace equation with multiple critical nonlinearities, Journal de Mathématiques Pures et Appliquées, 91 (2009), pp. 156–177.

[25] P. Guidotti, A new well-posed nonlinear nonlocal diffusion, Nonlinear Analysis: Theory, Methods & Applications, 72 (2010), pp. 4625–4637.

[26] P. Guidotti, A family of nonlinear diffusions connecting Perona-Malik to standard diffusion, Discrete & Continuous Dynamical Systems-Series S, 5 (2012).

[27] P. Guidotti, Y. Kim, and J. Lambers, Image Restoration with a New Class of Forward-Backward-Forward Diffusion Equations of Perona–Malik Type with Applications to Satellite Image Enhancement, SIAM Journal on Imaging Sciences, 6 (2013), pp. 1416–1444.

[28] P. Guidotti and K. Longo, Two enhanced fourth order diffusion models for image denoising, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 188–198.

[29] A. Hasanov, Inverse coefficient problems for monotone potential operators, Inverse Problems, 13 (1997), p. 1265.

[30] F. Hecht, New development in freefem++, Journal of Numerical Mathematics, 20 (2002), pp. 251–265.

[31] W. Hinterberger and O. Scherzer, Variational methods on the space of functions of bounded Hessian for convexification and denoising, Computing, 76 (2006), pp. 109–133.

[32] H. Houichet, A. Theljani, B. Rjaibi, and M. Moakher, A nonstandard higher-order variational model for speckle noise removal and thin-structure detection, Journal of Mathematical Study, 52 (2019), pp. 394–424.

[33] Z. Jin and X. Yang, A variational model to remove the multiplicative noise in ultrasound images, J. Math. Imaging Vis., 39 (2011), pp. 62–74.

[34] M. Kallel, M. Moakher, and A. Theljani, The Cauchy problem for a nonlinear elliptic equation: Nash-game approach and application to image inpainting, Inverse Problems and Imaging, 9 (2015), pp. 853–874.

[35] K. Krissian, R. Kikinis, C. F. Westin, and K. Vosburgh, Speckle-constrained filtering of ultrasound images, in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, 2005, pp. 547–552.

[36] P. Lindqvist, Notes on the p-Laplace Equation, University of Jyväskylä - Lectures notes, 2006.

[37] T. Loupas, Digital image processing for noise reduction in medical ultrasonics, PhD thesis, University of Edinburgh, UK, 1988.

[38] M. Lysaker, A. Lundervold, and X. C. Tai, Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Transactions on Image Processing, 12 (2003), pp. 1579–1590.

[39] M. Lysaker and X. C. Tai, Iterative image restoration combining total variation minimization and a second-order functional., International Journal of Computer Vision, 66 (2006), pp. 5–18.

[40] K. Papafitsoros and C.-B. Schönlieb, A combined first and second order variational approach for image reconstruction, Journal of Mathematical Imaging and Vision, 48 (2014), pp. 308–338.

[41] K. Papafitsoros and C.-B. Schönlieb, A combined first and second order variational approach for image reconstruction, Journal of mathematical imaging and vision, 48 (2014), pp. 308–338.

[42] V. S. Prasath, D. Vorotnikov, R. Pelapur, S. Jose, G. Seetharaman, and K. Palaniappan, Multiscale Tikhonov-total variation image restoration using spatially varying edge coherence exponent, IEEE Transactions on Image Processing, 24 (2015), pp. 5220–5235.

[43] O. Scherzer, Denoising with higher order derivatives of bounded variation and an application to parameter estimation, Computing, 60 (1998), pp. 1–27.

[44] S. Setzer and G. Steidl, Variational methods with higher-order derivatives in image processing, in Approximation Theory XII: San Antonio 2007, 2008.

[45] A. Theljani, Z. Belhachmi, M. Kallel, and M. Moakher, A multiscale fourth-order model for the image inpainting and low-dimensional sets recovery, Mathematical Methods in the Applied Sciences, 40 (2017), pp. 3637–3650.

[46] A. Theljani, Z. Belhachmi, and M. Moakher, High-order anisotropic diffusion operators in spaces of variable exponents and application to image inpainting and restoration problems, Nonlinear Analysis: Real World Applications, 47 (2019), pp. 251–271.

[47] D. Tschumperle and R. Deriche, Vector-valued image regularization with pdes: A common framework for different applications, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005), pp. 506–517.

[48] S. Wang, T.-Z. Huang, X.-L. Zhao, J.-J. Mei, and J. Huang, Speckle noise removal in ultrasound images by first-and second-order total variation, Numerical Algorithms, 78 (2018), pp. 513–533.

[49] J. Weickert, Theoretical foundations of anisotropic diffusion in image processing, in Theoretical Foundations of Computer Vision, W. Kropatsch, R. Klette, F. Solina, and R. Albrecht, eds., vol. 11 of Computing Supplement, Springer Vienna, 1996, pp. 221–236.

Liverpool Centre for Mathematics in Healthcare and Department of Mathematical Sciences, University of Liverpool, Liverpool L69 7ZL, UK

E-mail: `thaljanianis@gmail.com`