MOX-Report No. 40/2018

Mesh adaptation-aided image segmentation

Chiappa, A.S.; Micheletti, S.; Peli, R.; Perotto, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Mesh adaptation-aided image segmentation

Alberto Silvio Chiappa*, Stefano Micheletti#,
Riccardo Peli* and Simona Perotto#

June 20, 2018

MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica, Politecnico di Milano
Piazza L. da Vinci 32, I-20133 Milano, Italy
{stefano.micheletti,simona.perotto}@polimi.it

* Politecnico di Milano
Piazza L. da Vinci 32, I-20133 Milano, Italy
{alberto2.chiappa,riccardo.peli}@mail.polimi.it

2010 Mathematics Subject Classification: 49M25, 68U10, 65N30, 65N50, 65K10, 74G65
Keywords: image segmentation, variational model, anisotropic mesh adaptation, a posteriori error estimator

Abstract

We focus on a variational approach to image segmentation based on the Ambrosio-Tortorelli functional. To make the procedure more effective with respect to standard algorithms, we combine the functional minimization with the the employment of an optimal discretization. More precisely, we perform a finite element approximation of the Ambrosio-Tortorelli functional on a triangular adapted mesh able to follow exactly the contours present in the images, in the spirit of a *mesh adaptation-aided* image segmentation. This challenging goal is reached via a rigorous *a posteriori* error analysis enriched with anisotropic information. The benefits due to the proposed algorithm are evident both in terms of increased resolution in the edge detection and in a considerable reduction of the computational costs, as confirmed by an extensive numerical investigation.

1 Introduction

Image segmentation is central to image processing. Many practical applications require the identification of objects and boundaries (i.e., lines and curves, collectively referred to as segments), for instance X-ray based inspection techniques,

satellite imagery, medical diagnostics, texture and facial recognition, trajectory planning [28, 17, 7, 38, 8]. Due to the importance of devising fast and reliable algorithms to accomplish segmentation, a lot of research has been done in such a direction, leading to solutions that tackle the problem from different perspectives.

A first category comprises thresholding, clustering, edge-detection, and region-growing methods, among the others [27, 33, 35, 39].

A second class is represented by partial differential equation (PDE)-based approaches, such as the methods based on curve propagation (see, e.g., [11]). Among these, we cite the widely employed level set method [37]. The central idea here is to evolve a surface, ϕ, instead of a front, using a suitable signed function. The front is then defined implicitly as the zero level set, $\phi = 0$.

A third category is based on variational methods. The segmentation is here performed by minimizing a suitable energy functional, consisting of a data fitting contribution plus extra-regularization terms. One of the most known approaches is the one proposed in [29], based on the functional

$$I(u, E) = \int_\Omega (u - f)^2 \, d\Omega + \beta \int_{\Omega \setminus E} |\nabla u|^2 \, d\Omega + \gamma \, \mathcal{H}(E), \tag{1}$$

where $\Omega \subset \mathbb{R}^2$ represents the computational domain associated with the original image $f \in L^\infty(\Omega)$ (the true gray-level image), E is a closed one-dimensional (1D) subset of Ω of Hausdorff measure $\mathcal{H}(E)$, defining the separating edge, $u \in H^1(\Omega \setminus E)$ is the variational approximation to f, and β, γ are positive parameters to be tuned. The first term enforces that the approximation u be close to the original image f in the L^2-norm, while the second and third terms add regularization by penalizing the gradient of u and the length of E, respectively. In practice, this should ensure that the approximation u be quite flat in $\Omega \setminus E$, and that the edge set be as short as possible.
This mimimization problem, however, is very hard to be numerically tackled, mainly because of the presence of the low dimensional feature E, involving an unknown 1D entity embedded in a 2D framework.
In order to overcome this drawback, one can resort to the approximation introduced in [2, 3] by L. Ambrosio and V.M. Tortorelli, relying on the families of functionals $\{I_\epsilon\}_\epsilon$, depending on the positive parameter ϵ, defined by

$$I_\epsilon(u, v) = \int_\Omega (u-f)^2 \, d\Omega + \beta \int_\Omega (v^2+\eta) |\nabla u|^2 \, d\Omega + \gamma \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{1}{4\epsilon}(v-1)^2 \right) d\Omega, \tag{2}$$

with $\eta = \mathcal{O}(\epsilon^2)$ a positive parameter. The new function v, taking values in $[0, 1]$, plays the role of an approximate indicator of the set E, whose "thickness" has order ϵ. Moreover, both u and v are defined on the whole domain Ω, and no topological optimization of the geometric entity E is involved. The term $v^2 \, |\nabla u|^2$ forces v to get close to zero in the neighborhood of an edge, where a sharp variation of the image occurs. On the other hand, the last integral forces v

to be close to 1 almost everywhere. Overall, we expect that v be approximately constant with value 1, except across the edge, where it undergoes a steep drop to zero on a region on the order of ϵ. From a theoretical viewpoint, it can be proved that the sequence of functionals $\{I_\epsilon\}_\epsilon$ is Γ-convergent to the functional in (1), for $\epsilon \to 0$ [2, 3].

In this paper, we refer to a variational approach based on the functional in (2). In particular, we take advantage of the expertise gained in the modeling of crack propagation in brittle materials in [4, 5, 6]. Here, the crack evolution is detected by minimizing functional I_ϵ after neglecting the mismatch contribution, the two other terms representing the elastic and the fictitious crack energy, respectively. The phase field v is employed in both the contexts to track a contour (the crack rather than the image boundary), with the crucial difference that crack propagation is framed in a (quasi-static) evolving setting while image segmentation is a thoroughly stationary phenomenon. The original contribution in [4, 5, 6] has been to improve the efficiency of the minimization of the energy functional via a finite element discretization properly enriched with an anisotropic mesh adaptation procedure.

In the current work, this coupled minimization-mesh adaptation procedure is properly modified for image segmentation. Despite finite differences represent the most straightforward method to settle segmentation in a discrete context due to the intrinsic pixel structure of images, few contributions dealing with a finite element approximation can be found in the literature [9, 15, 20, 22, 41]. In particular, in [9] the author considers an algorithm that has some points in common with the one here presented, especially as far as the minimization part is concerned. Nevertheless, despite mentioning the possibility to implement some adaptation techniques, no suggestion is provided as to do this in practice, and a mesh coinciding with the pixel grid is used. This choice makes the computational cost of the algorithm very expensive for even slightly complex images. For this reason, we focus on mesh adaptation, as it both increases the resolution of the edges and greatly reduces the computational cost, in the spirit of a *mesh adaptation-aided* image segmentation. In the context of finite differences, in [16], the authors consider generalized relaxation methods coupled with multigrid linear solvers applied to the Euler-Lagrange equations of the Ambrosio-Tortorelli model.

To the best of our knowledge, no other paper in the literature has addressed anisotropic mesh adaptation in the context of image segmentation based on the Ambrosio-Tortorelli variational framework.

The paper is organized as follows. In Section 2, we focus on the minimization phase, tackled both in a continuous and in a discrete setting. Section 3 introduces the reference anisotropic framework, and provides the mathematical tool to drive the mesh adaptation procedure, i.e., an anisotropic *a posteriori* bound to control the functional error. We set the combined optimize-then-adapt algorithm in Section 4, after providing the practical rule to derive an optimal anisotropic metric from the *a posteriori* estimator. In Section 5, we assess the

proposed algorithm on synthetic as well as real images, by also verifying the best performances of the anisotropic approach with respect to standard techniques based on isotropic adapted and structured fixed meshes. Conclusions and perspectives are drawn in the last section. Finally, the explicit proof of the *a posteriori* analysis is provided in Appendix A for the sake of completeness.

2 The minimization process

The functional I_ϵ is clearly not jointly convex because of the term $v^2\,|\nabla u|^2$ so that the existence of a unique minimum cannot be trivially guaranteed. On the other hand, I_ϵ is convex with respect to both u and v, separately, which suggests proceeding with an alternating minimization, i.e., we fix alternately one of the two unknown functions and solve the minimization problem with respect to the other one. As a consequence, thanks to the Gâteaux differentiability of functional I_ϵ in $H^1(\Omega)\times(H^1(\Omega)\cap L^\infty(\Omega))$, we compute the two partial derivatives at (u,v) in the direction ϕ and ψ, respectively,

$$\begin{aligned} I_{\epsilon,u}(u,v;\phi) &= 2\int_\Omega (u-f)\,\phi\,d\Omega + 2\beta\int_\Omega (v^2+\eta)\nabla u\cdot\nabla\phi\,d\Omega := 2\,a(v;u,\phi),\\ I_{\epsilon,v}(u,v;\psi) &= 2\beta\int_\Omega |\nabla u|^2\,v\psi\,d\Omega + \frac{\gamma}{2\epsilon}\int_\Omega (v-1)\psi\,d\Omega + 2\gamma\epsilon\int_\Omega \nabla v\cdot\nabla\psi\,d\Omega\\ &\quad := 2\,b(u;v,\psi). \end{aligned} \tag{3}$$

Thus, minimization is obtained by setting

$$I'_\epsilon(u,v;\phi,\psi) := 2\,(a(v;u,\phi)+b(u;v,\psi)) = 0 \quad \forall(\phi,\psi)\in H^1(\Omega)\times(H^1(\Omega)\cap L^\infty(\Omega)). \tag{4}$$

Before setting the discrete couterpart, we highlight an important property of v which supports the intepretation of such a function as a smoothed indicator of the edges of f. Indeed, it holds:

Proposition 2.1 *If $(u,v)\in H^1(\Omega)\times(H^1(\Omega)\cap L^\infty(\Omega))$ is a critical point of $I_\epsilon(\cdot,\cdot)$, then $0\le v\le 1$ a.e. in Ω.*

For a proof of this result, we can adopt the same arguments as in [4, Proposition 2.3] and in [10, Proposition 1.3].

Remark 2.1 *An alternative proof of Proposition 2.1 can be obtained. Assume by contradiction the existence of a set, $\Omega_1\subset\Omega$, with positive measure, where $v>1$, so that $v=1+k$ in Ω_1, with k a positive function. It is readily checked that*

$$\hat v = \begin{cases} v & \text{in} \quad \Omega\setminus\overline{\Omega}_1\\ 1-k & \text{in} \quad \Omega_1 \end{cases}$$

still belongs to $H^1(\Omega)\cap L^\infty(\Omega)$. On comparing $I_\epsilon(u,\hat v)$ with $I_\epsilon(u,v)$, it holds that $I_\epsilon(u,\hat v) < I_\epsilon(u,v)$, since $\hat v^2 < v^2$, all the other terms being unchanged,

which contradicts that (u, v) is a critical point of $I_\epsilon(\cdot,\cdot)$. A similar argument can be employed to show that $v \geq 0$.

We now move to the discrete setting, by introducing a family $\{\mathcal{T}_h\}_{h>0}$ of triangular conformal meshes of the domain $\overline{\Omega}$ [13]. We associate with $\mathcal{T}_h$ the space

$$V_h = \{v \in C^0(\overline{\Omega}) : v|_K \in \mathbb{P}_1,\ \forall K \in \mathcal{T}_h\}, \tag{5}$$

of continuous piecewise linear finite elements.
We denote by $I_{\epsilon,h}(\cdot,\cdot)$ the discrete counterpart of $I_\epsilon(\cdot,\cdot)$, given by

$$\begin{aligned} I_{\epsilon,h}(u_h, v_h) &= \int_\Omega (u_h - f)^2\, d\Omega + \beta \int_\Omega (P_h(v_h^2) + \eta)|\nabla u_h|^2\, d\Omega \\ &+ \gamma \int_\Omega \left(\epsilon |\nabla v_h|^2 + \frac{1}{4\epsilon} P_h((v_h - 1)^2) \right) d\Omega, \end{aligned} \tag{6}$$

for any $(u_h, v_h) \in V_h^2$, where $P_h : C^0(\overline{\Omega}) \to V_h$ is the Lagrangian interpolant onto the space V_h [13]. Operator P_h is here introduced to guarantee the discrete analogue of the maximum principle in Proposition 2.1.

With a view to the minimization of $I_{\epsilon,h}(u_h, v_h)$, we are led to define the derivatives

$$\begin{aligned} I_{\epsilon,h,u}(u_h, v_h) &= 2 \int_\Omega (u_h - f)\, \phi_h\, d\Omega + 2\beta \int_\Omega (P_h(v_h^2) + \eta) \nabla u_h \cdot \nabla \phi_h\, d\Omega \\ &=: 2\, a_h(v_h; u_h, \phi_h) \\ I_{\epsilon,h,v}(u_h, v_h) &= 2\beta \int_\Omega |\nabla u_h|^2 P_h(v_h \psi_h)\, d\Omega + \frac{\gamma}{2\epsilon} \int_\Omega P_h((v_h - 1)\, \psi_h)\, d\Omega \\ &+ 2\gamma\epsilon \int_\Omega \nabla v_h \cdot \nabla \psi_h\, d\Omega =: 2\, b_h(u_h; v_h, \psi_h), \end{aligned} \tag{7}$$

for any $(\phi_h, \psi_h) \in V_h^2$. Analogously to [4, Proposition 2.5], the following statement can be proved.

Proposition 2.2 *If $(u_h, v_h) \in V_h^2$ is a critical point of $I_{\epsilon,h}(\cdot,\cdot)$, i.e., for any $(\phi_h, \psi_h) \in V_h^2$, $a_h(v_h; u_h, \phi_h) + b_h(u_h; v_h, \psi_h) = 0$, then $0 \leq v_h \leq 1$ in Ω.*

3 Anisotropic error estimator

This section is devoted to introducing the mathematical tool adopted to drive a sharp segmentation procedure of the true image. For this purpose, we first lay the anisotropic background according to [18, 31]. In particular, we exploit the spectral properties of the affine map, $T_K : \hat{K} \to K$, from the equilateral reference triangle, $\hat{K}$, inscribed in the unit circle, to the generic element, K, of $\mathcal{T}_h$, defined by

$$T_K(\hat{\mathbf{x}}) = M_K \hat{\mathbf{x}} + \mathbf{t}_K, \tag{8}$$

with $M_K \in \mathbb{R}^{2\times 2}$, $\mathbf{t}_K \in \mathbb{R}^2$, $\mathbf{x} = (x_1, x_2)^T \in K$, $\hat{\mathbf{x}} \in \hat{K}$.

With this aim, we first decompose M_K as $M_K = B_K Z_K$ with B_K, $Z_K \in \mathbb{R}^{2\times 2}$, a symmetric positive-definite and an orthogonal matrix, respectively. The two matrices, providing the so-called polar decomposition, deform and rotate $\hat{K}$ into K and the circle into an ellipse, $\mathcal{E}_K$, circumscribed to K. Matrix B_K is further decomposed as $B_K = R_K^T \Lambda_K R_K$, with $R_K^T = [\mathbf{r}_{1,K}, \mathbf{r}_{2,K}]$ and $\Lambda_K = \mathrm{diag}(\lambda_{1,K}, \lambda_{2,K})$, where $R_K \in \mathbb{R}^{2\times 2}$ is the eigenvector matrix, and $\Lambda_K \in \mathbb{R}^{2\times 2}$ collects the eigenvalues of B_K, with $\lambda_{1,K} \geq \lambda_{2,K}$. The geometric interpretation of these quantities identifies the eigenvectors $\mathbf{r}_{i,K}$ with the direction of the two semi-axes of $\mathcal{E}_K$, while the eigenvalues $\lambda_{i,K}$ meausure the corresponding length. The deformation of K is quantified by the aspect ratio, $s_K = \lambda_{1,K}/\lambda_{2,K} \geq 1$.

Next ingredient of the anisotropic setting is represented by suitable anisotropic estimates for the interpolation error associated with the quasi-interpolant Clément operator, $Q_h : L^2(\Omega) \to V_h$ [14], referring to [23, 24, 30] for all the details.

Lemma 3.1 *Let $w \in H^1(\Omega)$. If the cardinality $\#\Delta_K < \mathcal{N}$, for some $\mathcal{N} \in \mathbb{N}^+$, and if $\mathrm{diam}(T_K^{-1}(\Delta_K)) < C_\Delta \simeq \mathcal{O}(1)$, where $\Delta_K = \{T \in \mathcal{T}_h : T \cap K \neq \emptyset\}$, then there exist constants $C_s = C_s(\mathcal{N}, C_\Delta)$, with $s = 0, 1, 2$, such that, for any $K \in \mathcal{T}_h$, we have*

$$||w - Q_h(w)||_{H^r(K)} < C_r \left(\frac{1}{\lambda_{2,K}}\right)^r \left[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(w) \mathbf{r}_{i,K})\right]^{1/2}, \quad (9)$$

with $r = 0, 1$, and

$$||w - Q_h(w)||_{L^2(\partial K)} < C_2 \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2} \left[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(w) \mathbf{r}_{i,K})\right]^{1/2}, \quad (10)$$

where $h_K = \mathrm{diam}(K)$ and

$$G_{\Delta_K}(w) = \sum_{T\in\Delta_K} \begin{bmatrix} \displaystyle\int_T \left(\frac{\partial w}{\partial x_1}\right)^2 dT & \displaystyle\int_T \frac{\partial w}{\partial x_1}\frac{\partial w}{\partial x_2}\, dT \\ \displaystyle\int_T \frac{\partial w}{\partial x_1}\frac{\partial w}{\partial x_2}\, dT & \displaystyle\int_T \left(\frac{\partial w}{\partial x_2}\right)^2 dT \end{bmatrix} \quad (11)$$

is a positive semidefinite matrix.

Finally, we recall an equivalence result between the $H_1(\Delta_K)$-seminorm and its anisotropic analogue.

Lemma 3.2 *Let $w \in H^1(\Omega)$ and β_1, $\beta_2 > 0$. Then it holds that, for any $K \in \mathcal{T}_h$,*

$$\min\{\beta_1, \beta_2\} \leq \frac{\beta_1(\mathbf{r}_{1,K}^T G_{\Delta_K}(w)\mathbf{r}_{1,K}) + \beta_2(\mathbf{r}_{2,K}^T G_{\Delta_K}(w)\mathbf{r}_{2,K})}{|w|^2_{H^1(\Delta_K)}} \leq \max\{\beta_1, \beta_2\}. \quad (12)$$

We are now ready to furnish the main theoretical statement of this work, supporting the *a posteriori* error estimator which will drive mesh adaptation.

Proposition 3.1 *Let $(u_h, v_h) \in V_h^2$ be a critical point of $I_{\epsilon,h}(\cdot,\cdot)$. Then, it holds*

$$|I'_\epsilon(u_h, v_h; \phi, \psi)| \leq C \sum_{K\in\mathcal{T}_h} \left\{\rho_K^A(u_h, v_h)\, \omega_K(\phi) + \rho_K^B(u_h, v_h)\, \omega_K(\psi)\right\} \forall \phi, \psi \in H^1(\Omega), \tag{13}$$

where $I'_\epsilon(\cdot,\cdot;\cdot,\cdot)$ is defined as in (4), $C = C(\mathcal{N}, C_\Delta)$, while

$$\rho_K^A(u_h, v_h) = \|u_h - f\|_{L^2(K)} + 2\beta \left\|v_h(\nabla v_h \cdot \nabla u_h)\right\|_{L^2(K)} + \frac{\beta}{2} \|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)} \times \\ \|v_h^2 + \eta\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{\frac{1}{2}} + \frac{1}{\lambda_{2,K}} \|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)} \|\nabla u_h\|_{L^2(K)}, \tag{14}$$

$$\rho_K^B(u_h, v_h) = \left\| (\beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon}) v_h - \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} + \frac{\gamma\varepsilon}{2} \|[\![\nabla v_h]\!]\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}}\right)^{1/2} \\ + \frac{h_K^2}{\lambda_{2,K}} \left\| \beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} |v_h|_{W^{1,\infty}(K)}, \tag{15}$$

$$\omega_K(z) = \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(z) \mathbf{r}_{i,K}) \Big]^{1/2} \qquad \forall z \in H^1(\Omega), \tag{16}$$

where

$$[\![w_h]\!] = \begin{cases} \big|[\nabla w_h \cdot \mathbf{n}]\big| & \text{on } \mathcal{E}_h \\ \big|\nabla w_h \cdot \mathbf{n}\big| & \text{on } \mathcal{E}_h \cap \partial\Omega \end{cases} \tag{17}$$

denotes the absolute value of the jump of the normal derivative, with $\mathbf{n}$ the unit normal vector to the generic edge in the skeleton $\mathcal{E}_h$ of $\mathcal{T}_h$.

The proof of this proposition is a straightforward variant of [4, Proposition 3.3]. For the sake of completeness, it is provided in Appendix A.

Estimate (13) holds for any choice of test functions $(\phi, \psi) \in H^1(\Omega) \times (H^1(\Omega) \cap L^\infty(\Omega))$. Following [4, Corollary 3.4], picking $\phi = u - u_h$ and $\psi = v - v_h$ provides the term $I'_\epsilon(u_h, v_h; \phi, \psi)$ with a physical meaning, i.e., proportional to the energy error, $I_\epsilon(u, v) - I_\epsilon(u_h, v_h)$, up to a third-order remainder term.

Finally, to commute estimate (13) into an effective error estimator, we still need to replace the weights $\omega_K(u - u_h)$ and $\omega_K(v - v_h)$ with computable quantities, i.e.,

$$\omega_K^R(z) = \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}^R(z) \mathbf{r}_{i,K}) \Big]^{1/2} \qquad \text{with} \qquad z = u - u_h, v - v_h, \tag{18}$$

where the recovered matrix $G^R_{\Delta_K}(z)$ is defined by

$$\big[G^R_{\Delta_K}(z)\big]_{ij} = \sum_{T\in\Delta_K} \int_T \Big(\mathcal{R}_i(z_h) - \frac{\partial z_h}{\partial x_i}\Big)\Big(\mathcal{R}_j(z_h) - \frac{\partial z_h}{\partial x_j}\Big)\, dT, \qquad \text{with} \qquad i,j = 1,2,$$

with $[\mathcal{R}_1(z_h), \mathcal{R}_2(z_h)]^T$ the Zienkiewicz–Zhu recovered gradient of z_h [42, 43]. In particular, we employ the area-weighted recipe in [36], following [4, 5, 6, 21, 32].

Thus, the actual anisotropic error estimator is provided by $\zeta = \sum_{K\in\mathcal{T}_h} \zeta_K$, where

$$\zeta_K = \rho^A_K(u_h, v_h)\, \omega^R_K(u - u_h) + \rho^B_K(u_h, v_h)\, \omega^R_K(v - v_h) \tag{19}$$

defines the local contribution.

4 An adaptive algorithm for segmentation

Goal of this section is to build an efficient technique for image segmentation by exploiting the strength of mesh adaptation in an anisotropic setting. Such an approach is expected to be computationally advantageous compared to one using a fixed or an isotropically adapted mesh. For this purpose, we exploit the metric-based procedure proposed in [4], according to which the optimal metric induced by the estimator η is provided by the following

Proposition 4.1 *Let* `TOL` *be a tolerance demanded by the user on* ζ*. Then, the optimal elementwise metric* $\tilde\lambda^{-2}_{1,K}\tilde{\mathbf{r}}_{1,K}\tilde{\mathbf{r}}^T_{1,K} + \tilde\lambda^{-2}_{2,K}\tilde{\mathbf{r}}_{2,K}\tilde{\mathbf{r}}^T_{2,K}$ *equidistributing the error and guaranteeing a minimum number of elements, is provided by*

$$\begin{aligned} &\tilde\lambda_{1,K} = \left(\frac{1}{|\widehat{K}|\,\sqrt{2}}\left(\frac{g_{1,K}}{g^2_{2,K}}\right)^{1/2}\frac{\texttt{TOL}}{\#\mathcal{T}_h}\right)^{1/3}, \quad \tilde\lambda_{2,K} = \left(\frac{1}{|\widehat{K}|\,\sqrt{2}}\left(\frac{g_{2,K}}{g^2_{1,K}}\right)^{1/2}\frac{\texttt{TOL}}{\#\mathcal{T}_h}\right)^{1/3}, \\ &\tilde{\mathbf{r}}_{1,K} = \boldsymbol{\gamma}_{2,K}, \qquad\qquad\qquad\qquad \tilde{\mathbf{r}}_{2,K} = \boldsymbol{\gamma}_{1,K}, \end{aligned} \tag{20}$$

where $\{\boldsymbol{\gamma}_{i,K}, g_{i,K}\}$*, for* $i = 1, 2$*, are the eigenvector-eigenvalue pairs of matrix*

$$\Gamma_K = \big[\overline{\rho^A_K}(u_h, v_h)\big]^2\, \overline{G^R_{\Delta_K}}(u_h) + \big[\overline{\rho^B_K}(v_h, u_h)\big]^2\, \overline{G^R_{\Delta_K}}(v_h)$$

with $g_{1,K} \geq g_{2,K} > 0$*,*

$$\overline{\rho^m_K}(u_h, v_h) = \frac{\rho^m_K(u_h, v_h)}{\big(|\widehat{K}|\lambda_{1,K}\lambda_{2,K}\big)^{1/2}}$$

denotes the scaled residuals, with $m = A, B$*, and where*

$$\overline{G^R_{\Delta_K}}(z_h) = G^R_{\Delta_K}(z_h)/(|\widehat{K}|\,\lambda_{1,K}\lambda_{2,K})$$

is the scaled recovered matrix, for $z_h = u_h,\, v_h$*.*

Quantities in (20) are computed iteratively, starting from a uniform initial mesh. At each iteration, a new mesh is built in order to match the optimal metric in a predictive way. This means that all the quantities on the right-hand side in (20) are computed on the current mesh, while on the left-hand side we have the predicted values. For all the details, we refer to [4].

The mesh adaptation is now combined with the alternate minimization of functional $I_{\epsilon,h}(u_h, v_h)$ by Algorithm 1. In particular, we pursue the *optimize-then-adapt* algorithm in [4], where the minimimazion and the adaptation are loosely coupled to contain the overall computational cost.

Algorithm 1 `MAIStER`: Mesh Adaptation for Image Segmentation and Edge Recovery

Input: f, ϵ, η, β, γ, `TOL`, `TOLs`, `TOLfp`, `Ns`, `Nfp`, `TOLth`;
Output: $u_h, v_h, v_h^{\text{th}}, \mathcal{T}_h^*$;
Set `ERRs` = 1+`TOLs`, `i`=0;
while `ERRs` > `TOLs` & `i` < `Ns` **do**
    Set `ERRfp` = 1+`TOLfp`, `j`=0, if `i`=0 set $v_h^{(0)} = 1$;
    while `ERRfp` > `TOLfp` & `j` < `Nfp` **do**
        $u_h^{(\mathtt{j})} = \arg\min_{z_h \in V_h} I_{\epsilon,h}(z_h, v_h^{(\mathtt{j})})$;
        $v_h^{(\mathtt{j}+1)} = \arg\min_{z_h \in V_h} I_{\epsilon,h}(u_h^{(\mathtt{j})}, z_h)$;
        `ERRfp` = $\|v_h^{(\mathtt{j}+1)} - v_h^{(\mathtt{j})}\|_{L^\infty(\Omega)}$;
        `j` ← `j` + 1;
    end while
    Build the adapted mesh $\mathcal{T}_h^{(\mathtt{i}+1)}$ matching the metric based on $u_h^{(\mathtt{j}-1)}$ and $v_h^{(\mathtt{j})}$;
    `ERRs` = $|\#\mathcal{T}_h^{(\mathtt{i}+1)} - \#\mathcal{T}_h^{(\mathtt{i})}|/\#\mathcal{T}_h^{(\mathtt{i})}$;
    Set $v_h^{(0)} = \Pi_{\mathtt{i}\to\mathtt{i}+1}(v_h^{(\mathtt{j})})$;
    `i` ← `i` + 1;
end while
Set $u_h = \Pi_{\mathtt{i}\to\mathtt{i}+1}(u_h^{(\mathtt{j}-1)})$, $v_h = v_h^{(0)}$, $v_h^{\text{th}} = (v_h <$ `TOLth`$)$, $\mathcal{T}_h^* = \mathcal{T}_h^{(\mathtt{i}+1)}$;

Among the inputs, we provide original image, f, the parameters characterizing functional $I_\epsilon(\cdot,\cdot)$, together with the numerical ones driving the accuracy and the termination of both the **while** loops. In more detail, tolerances `TOLs` and `TOLfp` fix the accuracy on the mesh adaptation and on the alternate fixed-point procedure, while `Ns` and `Nfp` determine the corresponding maximum number of iterations. The value `TOLth` sets the threshold for extracting the edge set from the indicator function. The convergence of the mesh adaptivity is checked by monitoring the variation of the number of elements. An understood input quantity is the initial mesh, $\mathcal{T}_h^{(0)}$, coinciding with the structured mesh attached to the pixels of the true image f. The output quantities are the approximate image,

u_h, the approximate indicator, v_h, the thresholded approximate indicator, v_h^{th}, and the final adapted mesh, $\mathcal{T}_h^*$.

Operator $\Pi_{\mathtt{n}\to\mathtt{n}+1}$ is employed to interpolate a finite element function associated with mesh $\mathcal{T}_h^{(\mathtt{n})}$ onto the new mesh $\mathcal{T}_h^{(\mathtt{n}+1)}$.

The minimizations in lines 7.-8. involve two convex quadratic functionals and are tackled by standard direct or iterative algebraic solvers.

Algorithm 1 has been implemented in `FreeFem++` [26], while mesh adaptation is realized through the built-in function `adaptmesh` fed by the optimal metric in (20).

5 Results

In this section, we apply Algorithm 1 to two types of images, starting from synthetic tests, where the edges are clearly detectable, and then moving on to real images, representing more challenging configurations.

5.1 Synthetic images

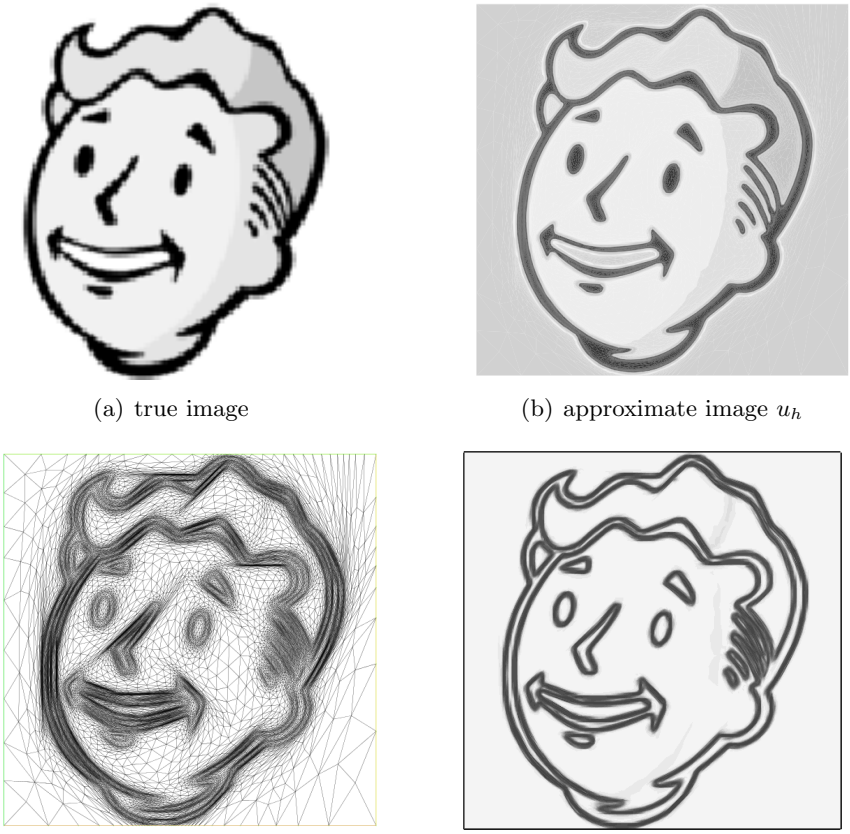
The focus of this section is to investigate the sensitivity of the Ambrosio-Tortorelli functional to the most relevant model parameters. For this purpose, we resort to three synthetic images to check the effect of changing one parameter at a time.

ϵ	$\max_K s_K$	# vertices	# elements
1	38.71	2654	5236
$5 \cdot 10^{-1}$	34.43	4455	8834
10^{-1}	140.91	11426	22752
$5 \cdot 10^{-2}$	381.39	16472	32855

Table 1: The Tintin test case: sensitivity to ϵ.

5.1.1 Sensitivity to ϵ

The first test case deals with a famous French cartoon (see Figure 1 (a)). Algorithm 1 is run with the data, $\epsilon = 5 \cdot 10^{-2}$, $\eta = 10^{-4}$, $\beta = 10^{-2}$, $\gamma = 4$, $\mathtt{TOL} = 50$, $\mathtt{TOLs} = 10^{-2}$, $\mathtt{TOLfp} = 5 \cdot 10^{-3}$, $\mathtt{Ns} = 10$, $\mathtt{Nfp} = 30$, $\mathtt{TOLth} = 10^{-1}$. After 7 iterations, the adaptive procedure stops. Figure 1 (b)-(d) collects some associated information. A cross comparison between the true image, consisting of 128×128 pixels, and the approximate image shows a very good match. Small detais, such as the sideburns and the dimple on the chin, are correctly detected. The final adapted mesh consists of 16472 nodes and 32855 elements and is characterized by a maximum aspect ratio of 381.39. Notice that, with fewer vertices, the approximate image is as sharp as the original one. The double edge contour of v_h is due to the capability of the error estimator to detect the gradient of



(a) true image

(b) approximate image u_h

(c) anisotropic adapted mesh

(d) approximate indicator v_h

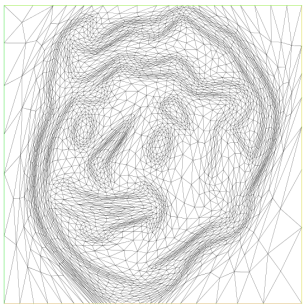
Figure 1: The Tintin test case.

the grayscale (for instance, the gray-black and black-white transition across the ear).

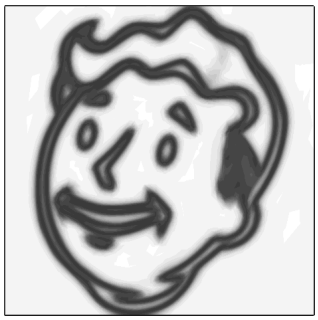
We now investigate the role played by ϵ, the other parameters being held fixed. In particular, we consider for ϵ the additional values 1, $5 \cdot 10^{-1}$, and 10^{-1}. Table 1 collects the associated maximum aspect ratio, number of vertices and elements, provided after 6 iterations of Algorithm 1 required to converge.

As expected, the number of vertices and elements increases as ϵ decreases, while the anisotropic shape of the triangles becomes more and more stretched.

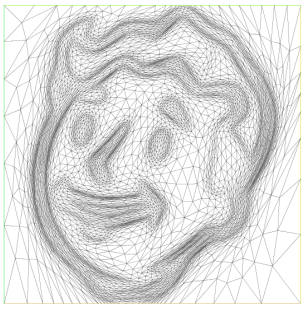
(a) $\epsilon = 1$: indicator v_h



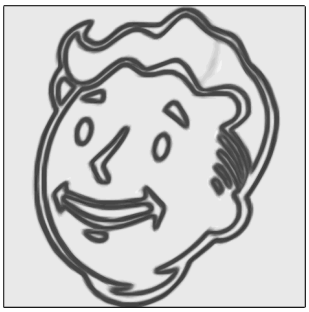
(b) $\epsilon = 1$: adapted mesh



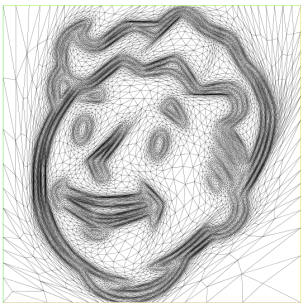
(c) $\epsilon = 5 \cdot 10^{-1}$: indicator v_h



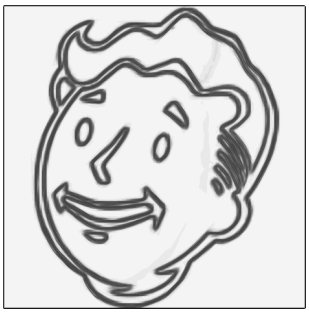
(d) $\epsilon = 5 \cdot 10^{-1}$: adapted mesh



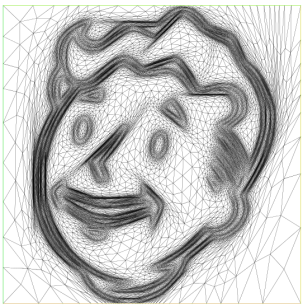
(e) $\epsilon = 10^{-1}$: indicator v_h



(f) $\epsilon = 10^{-1}$: adapted mesh



(g) $\epsilon = 5 \cdot 10^{-2}$: indicator v_h



(h) $\epsilon = 5 \cdot 10^{-2}$: adapted mesh

Figure 2: The Tintin test case: sensitivity to ϵ.

β	$\max_K s_K$	# vertices	# elements
1.5	97.78	19560	39077
1	66.01	16572	33097
$5 \cdot 10^{-1}$	107.31	12294	24553

Table 2: The pirate flag test case: sensitivity to β.

5.1.2 Sensitivity to β

The role of the parameter β in (1) is investigated on the synthetic image in Fig. 3 (top), showing a pirate flag. This parameter weights the regularization term based on the gradient of the approximate image u, outside of the edges. Consequently, we expect that for larger values of β, sharper contours are detected. For this purpose, we run Algorithm 1 with $\epsilon = 10^{-1}$, $\eta = 10^{-4}$, $\gamma = 4$, `TOL` $= 80$, `TOLs` $= 10^{-2}$, `TOLfp` $= 5 \cdot 10^{-3}$, `Ns` $= 10$, `Nfp` $= 50$, `TOLth` $= 10^{-1}$, and picking three values of β, i.e., 1.5, 1, and $5 \cdot 10^{-1}$. In Fig. 3 and in Tab. 2, we provide the qualitative and quantitative results of this analysis, respectively. In particular, Fig. 3 shows the indicator v_h and the corresponding anisotropic adapted mesh, while Tab. 2 collects the value of the maximum aspect ratio, and the number of vertices and triangles of the final adapted mesh. It is evident that, the thickness of the edge set decreases as β gets larger, and that the number of elements and vertices increases as well.

5.1.3 Sensitivity to γ

The numerical investigation is carried out on the image in Fig. 4, showing the Kármán vortex sheet on the cover of [40]. The images shows the streaklines exhibited by water flowing at 1.4 cm/s past a cylinder of diameter 1 cm at a Reynolds number equal to 140. The parameters to Algorithm 1 are $\epsilon = 10^{-1}$, $\eta = 10^{-4}$, $\beta = 5 \cdot 10^{-2}$, `TOL` $= 80$, `TOLs` $= 10^{-2}$, `TOLfp` $= 5 \cdot 10^{-3}$, `Ns` $= 10$, `Nfp` $= 50$, `TOLth` $= 10^{-1}$, and selecting three values of γ, i.e., 4, 1, and $2.5 \cdot 10^{-1}$. The parameter γ controls the regularization term of the length of the edges in (1). As a consequence, we expect that, for values of γ smaller and smaller, the outcome of the segmentation becomes blurred and the adapted meshes around the streaklines are thicker. These trends are confirmed by Fig. 4, collecting the

γ	$\max_K s_K$	# vertices	# elements
4	145.09	23455	46727
1	149.50	16246	32299
0.25	218.44	11109	22033

Table 3: The Kármán vortex sheet test case: sensitivity to γ.

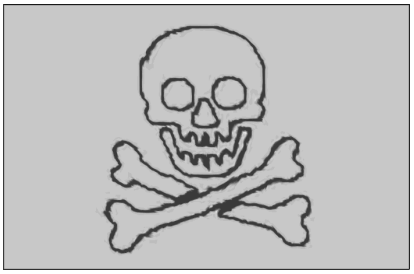
(a) true image

(b) $\beta = 1.5$: indicator v_h

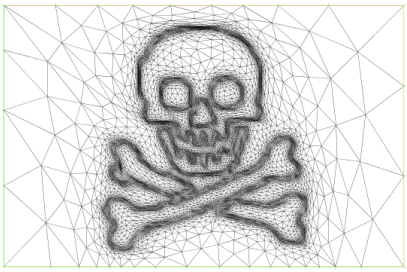
(c) $\beta = 1.5$: adapted mesh

(d) $\beta = 1$: indicator v_h

(e) $\beta = 1$: adapted mesh



(f) $\beta = 5 \cdot 10^{-1}$: indicator v_h

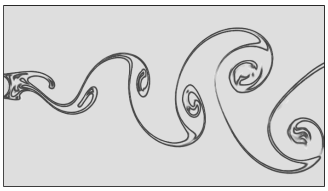


(g) $\beta = 5 \cdot 10^{-1}$: adapted mesh

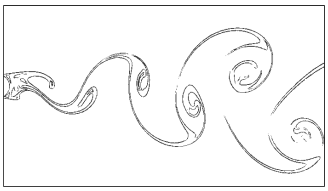
Figure 3: The pirate flag test case: sensitivity to β.

indicator v_h of the edge set (left), its thresholded version v_h^{th} (center), and the adapted mesh (right), and Tab. 3 gathering the same quantities as in Tables 1 and 2.

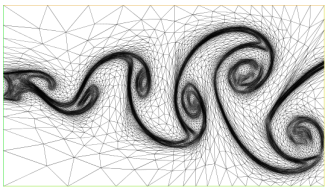
(a) true image



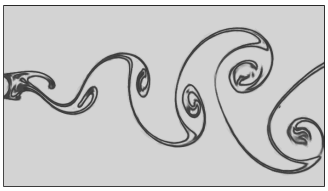
(b) $\gamma = 4$: indicator v_h



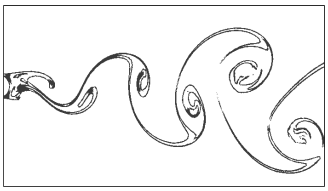
(c) $\gamma = 4$: thresholded indicator v_h^{th}



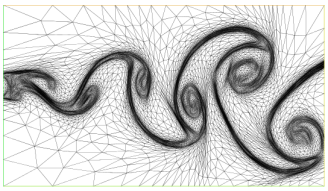
(d) $\gamma = 4$: adapted mesh



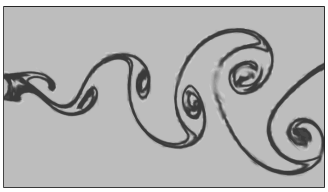
(e) $\gamma = 1$: indicator v_h



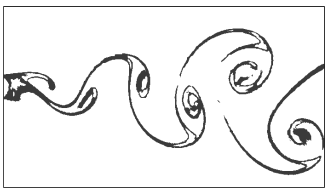
(f) $\gamma = 1$: thresholded indicator v_h^{th}



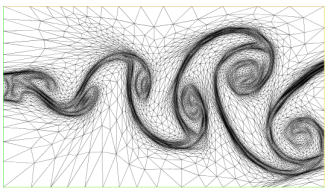
(g) $\gamma = 1$: adapted mesh



(h) $\gamma = 0.25$: indicator v_h



(i) $\gamma = 0.25$: thresholded indicator v_h^{th}



(j) $\gamma = 0.25$: adapted mesh

Figure 4: The Kármán vortex sheet test case: sensitivity to γ.

5.2 Real images

We now consider the segmentation of real images. We focus on medical images due to their strong impact on the health and social care. Moreover, the presence of both soft tissues and bones makes this analysis extremely challenging.

In particular, we pick the following images:

1. The MRI of a brain
   `https://it.m.wikipedia.org/wiki/File:Pkan-basal-ganglia-MRI.JPG`
2. The MRA of the circle of Willis
   `https://www.spandidos-publications.com/br/3/1/55`

3. The X-ray of a knee
   `https://www.howardluksmd.com/orthopedic-social-media/medial-joint-space-narrowing/`

4. The X-ray of a hand
   `https://it.wikipedia.org/wiki/File:Artrite_psoriasica_Rx_Mano_Sn.PNG`

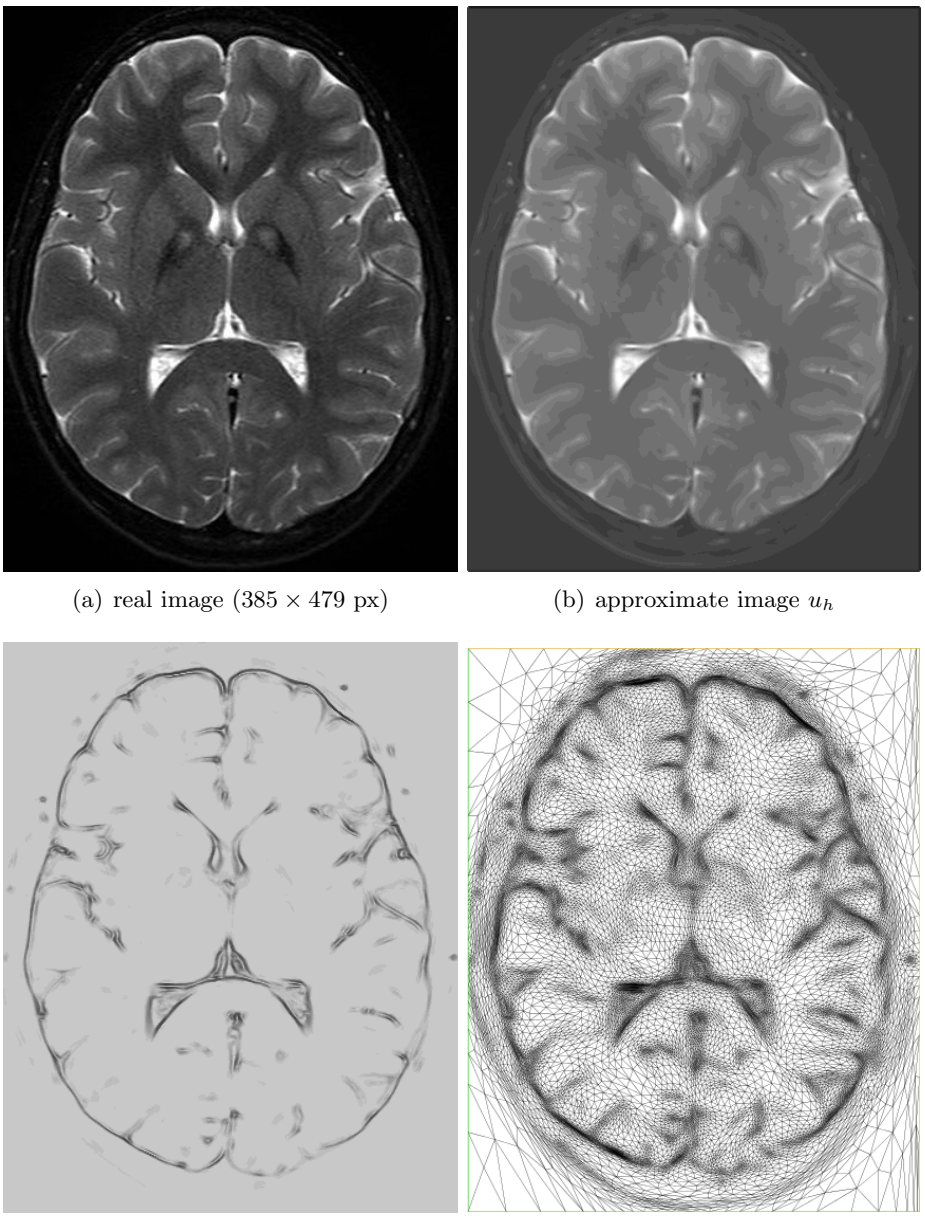
Figures 5-8 (top-left) show the original images along with the corresponding resolution in pixels (px). Table 4 collects the input parameters to Algorithm 1. For this choice of parameters, a fast convergence is always attained and, in all cases, the fixed point method takes a few iterations. In Table 5, we gather the main output values of Algorithm 1, namely, the maximun aspect ratio, the number of vertices and triangles of the last adapted mesh, the number of global iterations and the corresponding number of fixed point iterations. The quality of the segmented images, provided in Fig. 5-8 (top-right), is high, despite the relative small number of elements and vertices, as compared with the number of pixels in the original images. In more detail, in Fig. 5, the segmentation procedure identifies the plural gyri and sulci of the grey matter, together with the white matter, the ventricles and the corpus callosum. Figure 6 clearly displays the circle of Willis and the slender posterior cerebral artery. Figure 7 reveals the three main bones, femur, tibia and fibula, and some parts of muscles and soft tissues. In the last Fig. 8, the joint deformation associated with a psoriatic arthritis is detected, together with the carpals, metacarpals, and the three phalanges.

The indicator function v_h, for a threshold set to $\texttt{TOLth} = 0.1$, highlights essentially the high-contrast areas, as shown in Figures 5-8 (bottom-left).

Finally, concerning the adapted meshes, shown in Figures 5-8 (bottom-right), they all ensure a high quality segmentation, although being quite coarse, in particular for the knee and hand images, and mildly anisotropic.

In Table 6, we show the computational times, in seconds[1], demanded by Algorithm 1 for all the real images in Figures 5-8. We collect the execution time in the first column, while we distiguish between the time required by the adaptation (estimator and metric computation plus mesh generation) and by the minimization phase, in the second and third column, respectively. In particular, we provide explicitly the time for each fixed point iteration. We observe that the dominant contribution is the one due to the adaptation procedure, with the highest value at the first iteration when the computational mesh coincides with the stuctured triangual mesh associated with the original image.

[1]The computations have been run on a GenuineIntel Pentium(R) Dual-Core CPU E6300 2.80 GHz 4GB RAM desktop computer.



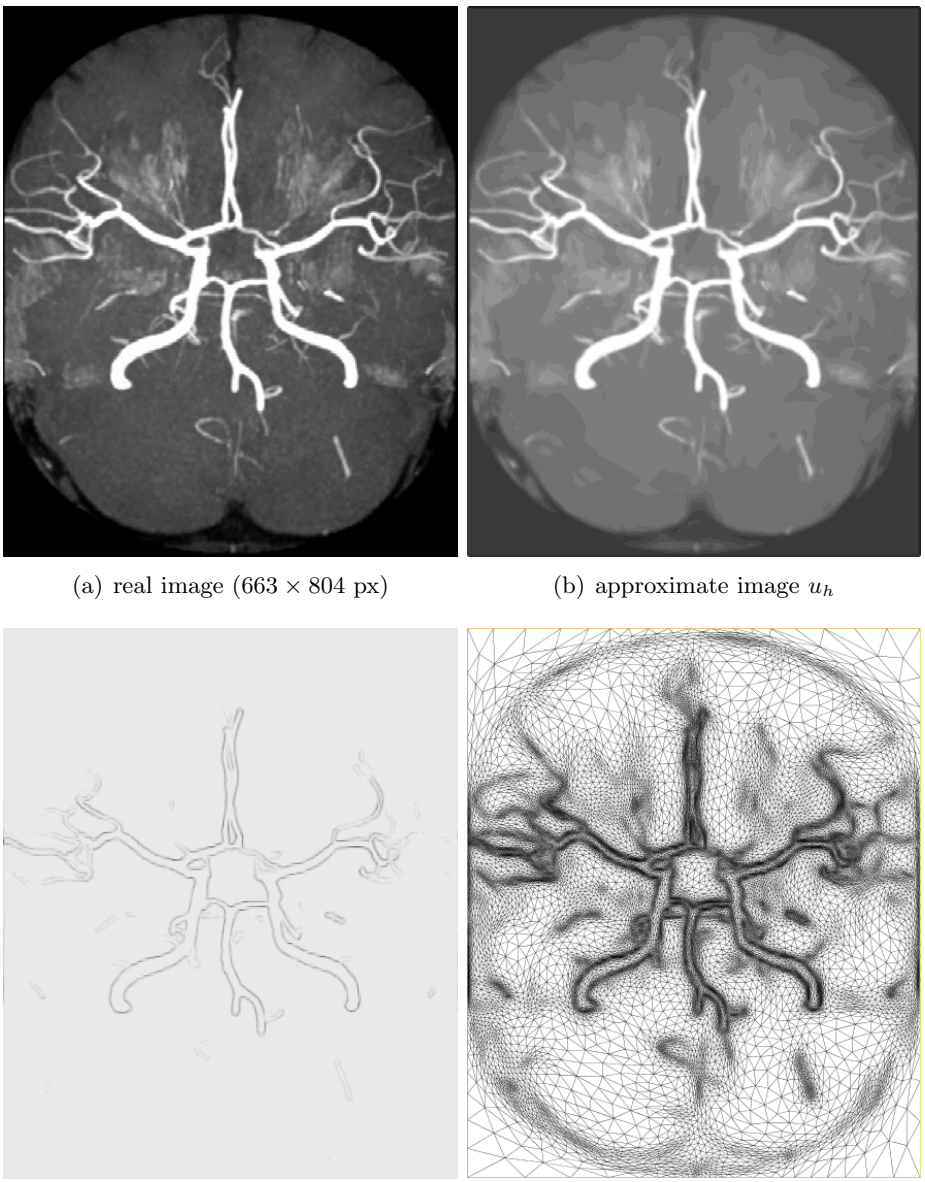
(a) real image (385×479 px)

(b) approximate image u_h

(c) indicator v_h

(d) adapted mesh

Figure 5: The brain test case.



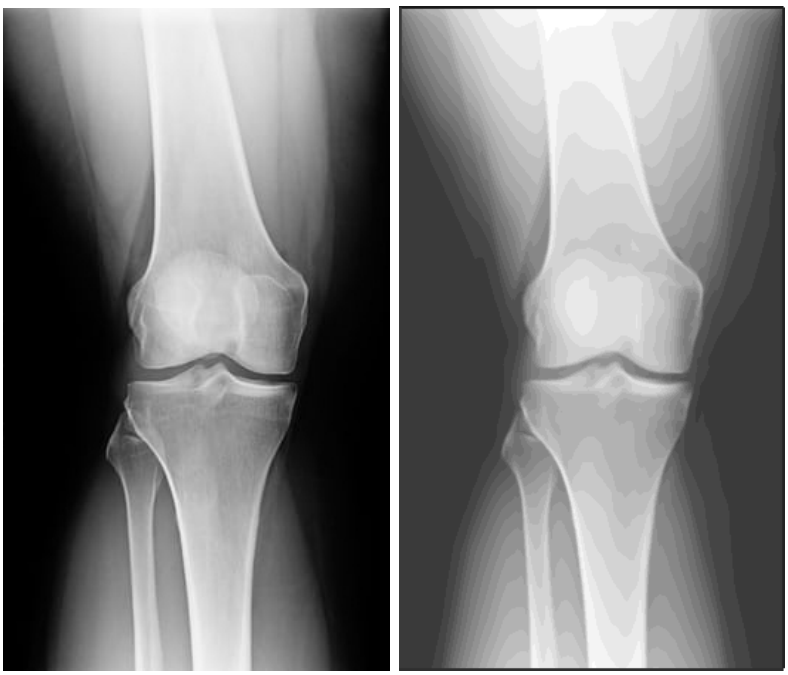
(a) real image (663×804 px)

(b) approximate image u_h

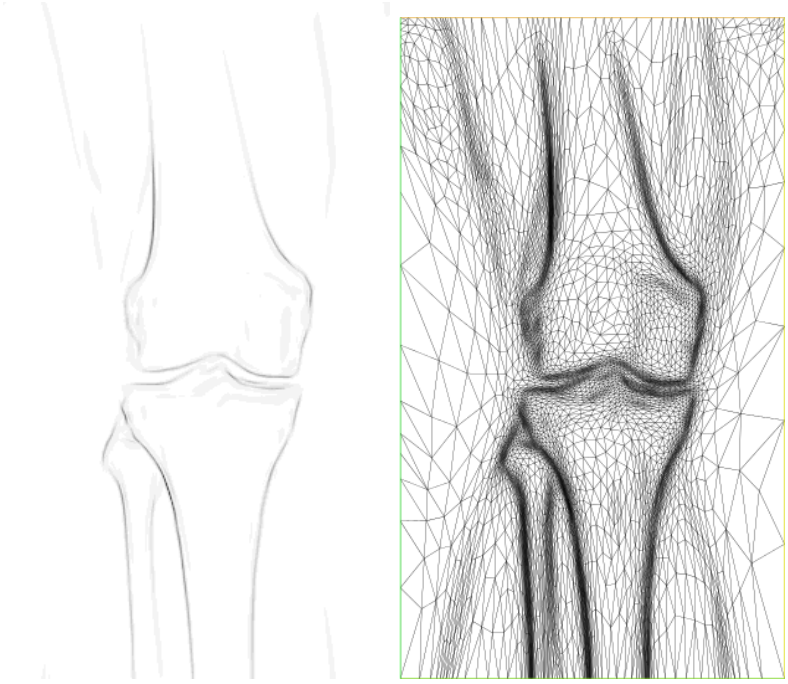
(c) indicator v_h

(d) adapted mesh

Figure 6: The circle of Willis test case.

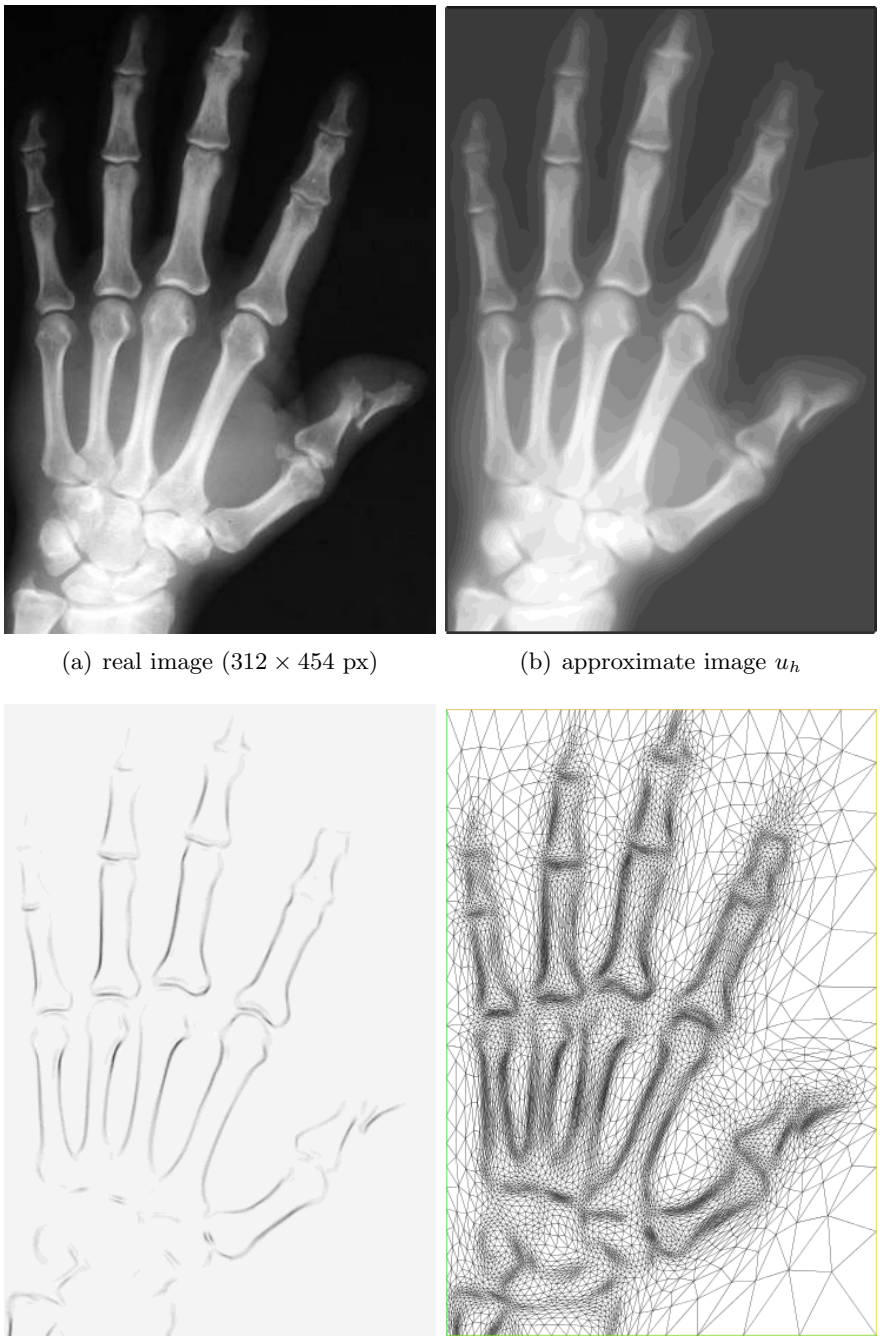


(a) real image (261 × 448 px) (b) approximate image u_h



(c) indicator v_h (d) adapted mesh

Figure 7: The knee test case.



(a) real image (312×454 px) (b) approximate image u_h

(c) indicator v_h (d) adapted mesh

Figure 8: The hand test case.

image	ϵ	β	γ	η	TOL	TOLs	TOLfp	Ns	Nfp
brain	10^{-3}	10^{-1}	$5 \cdot 10^{-1}$	10^{-6}	70	10^{-2}	$5 \cdot 10^{-3}$	10	40
Willis	10^{-3}	10^{-1}	5	10^{-6}	100	10^{-2}	$5 \cdot 10^{-3}$	10	40
knee	10^{-3}	$2 \cdot 10^{-1}$	10	10^{-6}	60	10^{-2}	$5 \cdot 10^{-3}$	10	40
hand	10^{-3}	10^{-1}	10	10^{-6}	60	10^{-2}	$5 \cdot 10^{-3}$	10	40

Table 4: Real images: input parameters to Algorithm 1.

image	$\max_K s_K$	# vertices	# elements	# ITs	# ITfp
brain	71.94	19726	39334	4	$4-8-9-25$
Willis	56.51	24845	49544	3	$2-4-4$
knee	82.01	5683	11175	4	$3-5-7-6$
hand	38.01	8592	17056	5	$2-3-2-2-2$

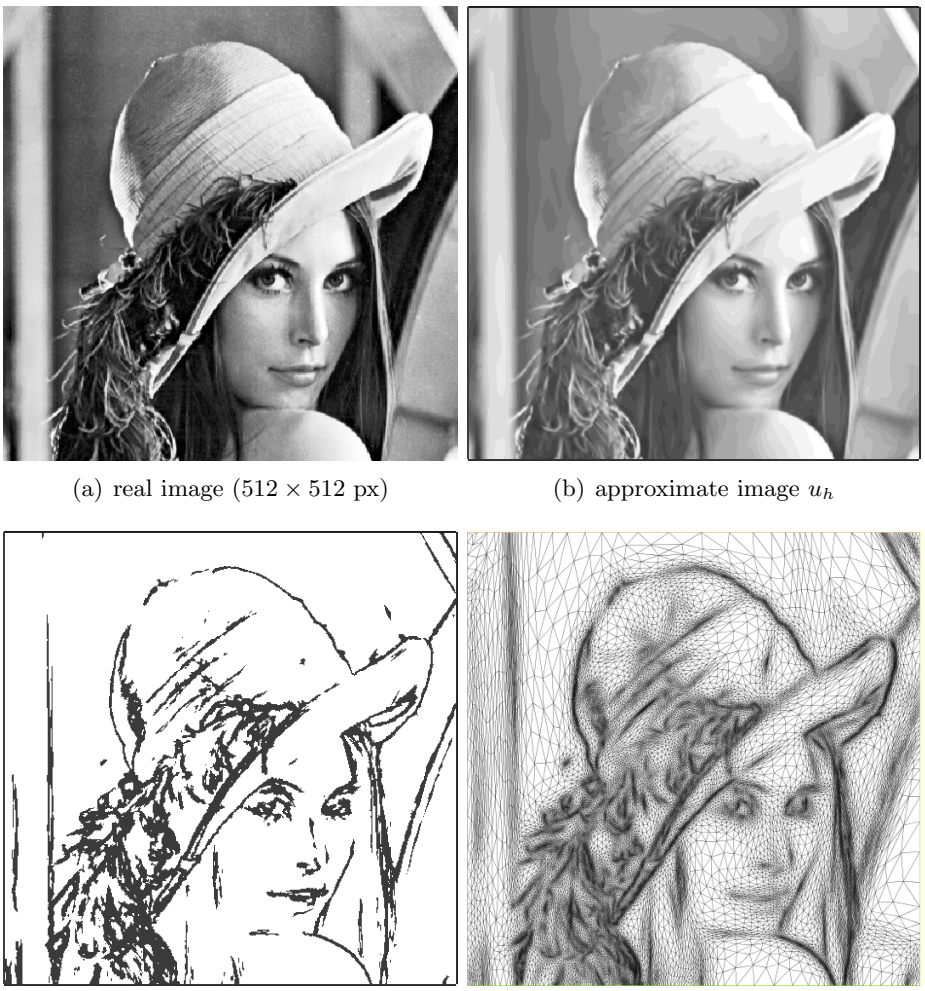
Table 5: Real images: output of Algorithm 1.

image	E-time	AD-time	FP-time
brain	170.74	$43.77-24.95-19.45-17.58$	$11.26-12.57-10.62-25.51$
Willis	102.78	$29.76-24.12-22.93$	$3.63-5.78-5.3$
knee	56.29	$23.56-7.66-5.56-4.64$	$4.87-2.31-2.24-1.54$
hand	79.90	$29.83-11.32-8.53-7.48-7.3$	$4.14-2.01-0.97-0.87-0.81$

Table 6: Real images: execution (E-time), adaptive (AD-time) and fixed point (FP-time) iteration times.

5.3 Comparison with a pixel mesh

We refer to the test case of Fig. 8 in [9], dealing with the well-known Lena image benchmark (see Fig. 9 (top-left)). We adopt the following parameters, $\epsilon = 10^{-1}$, $\eta = 10^{-2}$, $\beta = 2$, $\gamma = 10^{-2}$, to match those in [9], in addition to $\texttt{TOL} = 50$, $\texttt{TOLs} = 10^{-2}$, $\texttt{TOLfp} = 5 \cdot 10^{-3}$, $\texttt{Ns} = 10$, $\texttt{Nfp} = 50$, $\texttt{TOLth} = 8 \cdot 10^{-5}$. Figure 9 collects the outcomes of Algorithm 1 at convergence (after 7 iterations), represented by the approximation to the original image, the edge set, and the final adapted mesh, consisting of 67679 anisotropic triangles with a maximum aspect ratio equal to 75.58. The quality of the edge set is thoroughly comparable with the one in the reference paper, although the number of elements is considerably lower for the adaptive strategy.



(a) real image (512×512 px) (b) approximate image u_h

(c) thresholded indicator v_h^{th} (d) adapted mesh

Figure 9: The Lena test case.

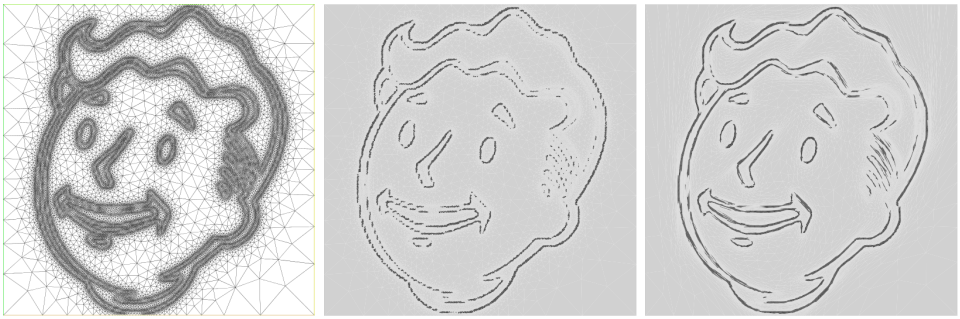
5.4 Comparison with the isotropic mesh adaptation

The advantages due to anisotropic mesh adaptation with respect to fixed or isotropically adapted meshes are well established [25, 1, 19]. For the sake of completeness, we run again some of the previous test cases in an isotropic setting to confirm these improvements.

In general, for the same accuracy `TOL`, the isotropic metric is obtained by enforcing $s_K = 1$, for any $K \in \mathcal{T}_h$, i.e., replacing the recipes in Proposition 4.1 by

$$\tilde{\lambda}_{1,K} = \tilde{\lambda}_{2,K} = \left(\frac{1}{|\widehat{K}|\sqrt{2}} \left(\frac{2}{g_{1,K} + g_{2,K}} \right)^{1/2} \frac{\texttt{TOL}}{\#\mathcal{T}_h} \right)^{1/3}, \quad \tilde{\mathbf{r}}_{1,K} = [1,0]^T, \quad \tilde{\mathbf{r}}_{2,K} = [0,1]^T. \tag{21}$$

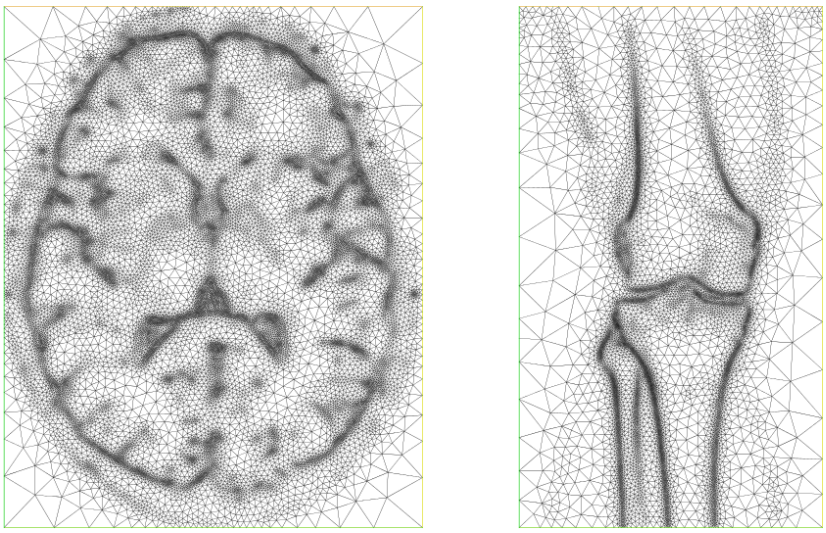
We have performed two series of checks. The results of the first check, provided in Fig. 10, refer to the test case in Section 5.1.1. The isotropic mesh in the left panel consists of 62023 triangles. Despite the large number of elements, the quality of the approximate thresholded indicator, with `TOLth` $= 0.1$ (center panel), is low compared with the corresponding function (right panel) computed on the anisotropic mesh in Fig. 1 (c). Actually, the main features of the image are captured but some of the details are lost and detected discontinuously.



(a) isotropic adapted mesh (b) isotropic thresholded indicator (c) anisotropic thresholded indicator

Figure 10: The Tintin test case: isotropic vs anisotropic mesh adaptation.

In the second check, we compare the isotropic and anisotropic procedures for the same accuracy `TOL`. We rerun the test cases in Fig. 5 and in Fig. 7, replacing formulas in (20) by the ones in (21). In both cases, the number of elements required by the isotropic approach is larger, being 48829 and 19760 in the brain and knee test case, respectively (see Fig. 11).



(a) brain test case (b) knee test case

Figure 11: Real images: isotropic adaptive mesh for the brain (left) and knee (right) test case.

6 Conclusions and perspectives

The extensive numerical investigation shows that the proposed adaptive method is reliable on both synthetic and real images. We succeed in effectively segmenting even complex images as those characterizing medical applications. This is obtained on quite coarse meshes in contrast to the original pixel representation, or to isotropically adapted grids. The additional value of Algorithm 1 is to provide a high quality compression tool, supporting both the approximate image as well as the edge set.

In perspective, we are extending the proposed procedure to 3D images, with particular emphasis on medical MRI/MRA data. This represents a crucial step with a view to computational fluid-dynamics simulations in the reconstructed geometries.

A Proof of Proposition 3.1

Since (u_h, v_h) is a critical point of $I_{\epsilon,h}(\cdot,\cdot)$, we have

$$a_h(v_h; u_h, \phi_h) = 0 \quad \forall \phi_h \in V_h, \quad b_h(u_h; v_h, \psi_h) = 0 \quad \forall \psi_h \in V_h. \tag{22}$$

Moreover, moving from (3) and using the triangular inequality, it holds that

$$|I'_\epsilon(u_h, v_h; \phi, \psi)| \leq 2(|a(v_h; u_h, \phi)| + |b(u_h; v_h, \psi)|), \tag{23}$$

for any pair $(\phi, \psi) \in H^1(\Omega) \times (H^1(\Omega) \cap L^\infty(\Omega))$. We now estimate the two terms above, separately.

Estimate of $|a(v_h; u_h, \phi)|$.

Thanks to (22) and to the linearity of $a(\cdot; \cdot, \phi)$ with respect to ϕ, for any $\phi \in H^1(\Omega), \phi_h \in V_h$, we have

$$|a(v_h; u_h, \phi)| \leq |a(v_h; u_h, \phi - \phi_h)| + |a(v_h; u_h, \phi_h) - a_h(v_h; u_h, \phi_h)|. \tag{24}$$

Let us deal with the first term on the right-hand side of (24). Concerning the first one and using $(3)_1$, we have

$$\begin{aligned}
&\big|a(v_h; u_h, \phi - \phi_h)\big| \\
&= \Big| \sum_{K \in \mathcal{T}_h} \Big\{ \int_K (u_h - f)(\phi - \phi_h)\, dK + \beta \int_K (v_h^2 + \eta) \nabla u_h \cdot \nabla(\phi - \phi_h)\, dK \Big\} \Big| \\
&= \Big| \sum_{K \in \mathcal{T}_h} \Big\{ \int_K (u_h - f)(\phi - \phi_h)\, dK - \beta \int_K 2 v_h (\nabla v_h \cdot \nabla u_h)(\phi - \phi_h)\, dK \\
&+ \beta \int_{\partial K} (v_h^2 + \eta) \nabla u_h \cdot \mathbf{n} (\phi - \phi_h)\, ds \Big\} \Big| \\
&\leq \sum_{K \in \mathcal{T}_h} \Big\{ \Big(\|u_h - f\|_{L^2(K)} + 2\beta \, \|v_h (\nabla v_h \cdot \nabla u_h)\|_{L^2(K)} \Big) \|\phi - \phi_h\|_{L^2(K)} \\
&+ \frac{\beta}{2} \|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)} \|v_h^2 + \eta\|_{L^2(\partial K)} \|\phi - \phi_h\|_{L^2(\partial K)} \Big\}
\end{aligned}$$

after splitting the integrals on the mesh elements, exploiting integration by parts, Hölder and Cauchy-Schwarz inequalities, and definition (17). Choosing $\phi_h = Q_h(\phi)$ and using results (9) with $s = 0$, and (10), we obtain

$$\begin{aligned}
\big|a(v_h; u_h, \phi - \phi_h)\big| \leq C \sum_{K \in \mathcal{T}_h} \Big\{ &\|u_h - f\|_{L^2(K)} + 2\beta \, \|v_h (\nabla v_h \cdot \nabla u_h)\|_{L^2(K)} \\
&+ \frac{\beta}{2} \|[\![\nabla u_h]\!]\|_{L^\infty(\partial K)} \|v_h^2 + \eta\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K} \lambda_{2,K}} \right)^{\frac{1}{2}} \Big\} \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(\phi) \mathbf{r}_{i,K}) \Big]^{\frac{1}{2}},
\end{aligned} \tag{25}$$

with $C = \max(C_0, C_2)$. Let us now deal with the second term on the right-hand side of (24). Using Hölder and Cauchy-Schwarz inequalities yields

$$\begin{aligned}
&|a(v_h; u_h, \phi_h) - a_h(v_h; u_h, \phi_h)| = \Big| \beta \int_\Omega \big[v_h^2 - P_h(v_h^2) \big] \nabla u_h \cdot \nabla \phi_h \, d\Omega \Big| \\
&\leq \beta \sum_{K \in \mathcal{T}_h} \Big\{ \|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)} \, \|\nabla u_h\|_{L^2(K)} \, \|\nabla \phi_h\|_{L^2(K)} \Big\}.
\end{aligned} \tag{26}$$

Then, thanks to (9) with $s = 1$, together with Lemma 3.2 with $\beta_1 = \lambda_{1,K}^2$, $\beta_2 = \lambda_{2,K}^2$, this becomes

$$\begin{aligned} & |a(v_h; u_h, \phi_h) - a_h(v_h; u_h, \phi_h)| \\ \leq\ & \beta \sum_{K \in \mathcal{T}_h} \Big\{ \Big(\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)} \, \|\nabla u_h\|_{L^2(K)} \Big) \\ & \Big(\|\nabla \phi_h - \nabla \phi\|_{L^2(K)} + \|\nabla \phi\|_{L^2(K)} \Big) \Big\} \\ \leq\ & C \sum_{K \in \mathcal{T}_h} \Big\{ \Big(\|v_h^2 - P_h(v_h^2)\|_{L^\infty(K)} \, \|\nabla u_h\|_{L^2(K)} \Big) \\ & \frac{1}{\lambda_{2,K}} \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(\phi) \mathbf{r}_{i,K}) \Big]^{1/2} \Big\}, \end{aligned} \tag{27}$$

with $C = \beta(C_1 + 1)$. The results obtained in (25) and (27), and definitions (14), (16) allow us to bound the first term on the right-hand side of (23), as

$$|a(v_h; u_h, \phi)| \leq C_A \sum_{K \in \mathcal{T}_h} \rho_K^A(u_h, v_h) \, \omega_K(\phi), \tag{28}$$

where $C_A = \max\big(\max(C_0, C_2), \beta(C_1 + 1) \big)$.

Estimate of $|b(u_h; v_h, \psi)|$.

Thanks to (22) and to the linearity of $b(\cdot; \cdot, \psi)$, for any $\psi \in H^1(\Omega), \psi_h \in V_h$, it holds

$$|b(u_h; v_h, \psi)| \leq |b(u_h; v_h, \psi - \psi_h)| + |b(u_h; v_h, \psi_h) - b_h(u_h; v_h, \psi_h)|. \tag{29}$$

We tackle the first term. Splitting the integrals over the mesh elements, integrating by parts, and thanks to the Cauchy-Schwarz inequality and definition (17), we obtain

$$\begin{aligned} & \big| b(u_h; v_h, \psi - \psi_h) \big| \\ = & \left| \sum_{K \in \mathcal{T}_h} \left\{ \int_K \left[\left((\beta |\nabla u_h|^2 + \frac{\gamma}{4\varepsilon}) v_h - \frac{\gamma}{4\varepsilon} \right) (\psi - \psi_h) + \gamma \varepsilon \nabla v_h \cdot \nabla (\psi - \psi_h) \right] dK \right\} \right| \\ \leq & \sum_{K \in \mathcal{T}_h} \left\{ \left\| (\beta |\nabla u_h|^2 + \frac{\gamma}{4\varepsilon}) v_h - \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} \|\psi - \psi_h\|_{L^2(K)} \right. \\ & \qquad \left. + \left| \gamma \varepsilon \int_{\partial K} (\psi - \psi_h) \nabla v_h \cdot \mathbf{n} \, ds \right| \right\} \\ \leq & \sum_{K \in \mathcal{T}_h} \left\{ \left\| (\beta |\nabla u_h|^2 + \frac{\gamma}{4\varepsilon}) v_h - \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} \|\psi - \psi_h\|_{L^2(K)} \right. \\ & \qquad \left. + \frac{\gamma \varepsilon}{2} \, \|[\![\nabla v_h]\!]\|_{L^2(\partial K)} \|\psi - \psi_h\|_{L^2(\partial K)} \right\}. \end{aligned} \tag{30}$$

We now choose $\psi_h = Q_h(\psi)$. Thanks to (9) for $s = 0$ and (10), it follows

$$
\begin{aligned}
&\big|b(u_h; v_h, \psi - \psi_h)\big| \\
&\leq C \sum_{K\in\mathcal{T}_h} \left\{ \left\| (\beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon}) v_h - \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} + \frac{\gamma\varepsilon}{2} \left\| [\![\nabla v_h]\!] \right\|_{L^2(\partial K)} \left(\frac{h_K}{\lambda_{1,K}\lambda_{2,K}} \right)^{1/2} \right\} \\
&\left[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(\psi) \mathbf{r}_{i,K}) \right]^{1/2},
\end{aligned}
\tag{31}
$$

the constant being $C = \max(C_0, C_2)$. The estimation of the second term on the right-hand side of (29) can be performed by combining the Cauchy-Schwarz inequality with the standard isotropic estimate for the L^2-norm of the interpolation error associated with P_h and the auxiliary result $|w_h\varphi_h|_{H^2(K)} \leq 2\,|w_h|_{W^{1,\infty}(K)}\,\|\nabla\varphi_h\|_{L^2(K)}$ for any $w_h, \varphi_h \in V_h$ and for any $K \in \mathcal{T}_h$, so that

$$
\begin{aligned}
&\big|b(u_h; v_h, \psi_h) - b_h(u_h; v_h, \psi_h)\big| \\
&= \Big| \int_\Omega (v_h\psi_h - P_h(v_h\psi_h)) \left(\beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right) d\Omega \Big| \\
&\leq \sum_{K\in\mathcal{T}_h} \left\{ \|v_h\psi_h - P_h(v_h\psi_h)\|_{L^2(K)} \left\| \beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} \right\} \\
&\leq C_I \sum_{K\in\mathcal{T}_h} \left\{ \left\| \beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} h_K^2\, |v_h\psi_h|_{H^2(K)} \right\} \\
&\leq 2C_I \sum_{K\in\mathcal{T}_h} \left\{ \left\| \beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} h_K^2\, |v_h|_{W^{1,\infty}(K)} \|\nabla\psi_h\|_{L^2(K)} \right\} \\
&\leq C \sum_{K\in\mathcal{T}_h} \left\{ \left\| \beta|\nabla u_h|^2 + \frac{\gamma}{4\varepsilon} \right\|_{L^2(K)} h_K^2\, |v_h|_{W^{1,\infty}(K)} i\times \right. \\
&\qquad \left. \frac{1}{\lambda_{2,K}} \Big[\sum_{i=1}^{2} \lambda_{i,K}^2 (\mathbf{r}_{i,K}^T G_{\Delta_K}(\psi) \mathbf{r}_{i,K}) \Big]^{1/2} \right\},
\end{aligned}
\tag{32}
$$

with $C = 2C_I(1 + C_1)$, C_I being the constant in the interpolation estimate, and where the last inequality follows, proceeding as in (26), on using (9) with $s = 1$ and Lemma 3.2 with $\beta_1 = \lambda_{1,K}^2$ and $\beta_2 = \lambda_{2,K}^2$.
Inequalities (31) and (32), together with definitions (15), (16), allow us to control the second term on the right-hand side of (23) as

$$
|b(u_h; v_h, \psi)| \leq C_B \sum_{K\in\mathcal{T}_h} \rho_K^B(u_h, v_h)\, \omega_K^B(\psi),
\tag{33}
$$

with $C_B = \max(\max(C_0, C_2), 2C_I(1 + C_1))$. Estimate (13) now follows in a straightforward way by combining estimates (28) and (33), where $C = \max(C_A, C_B)$.

Acknowledgments

The second and fourth authors acknowledge the research project GNCS-INdAM 2018 "Tecniche di Riduzione di Modello per le Applicazioni Mediche".

References

[1] D. Ait-Ali-Yahia, G. Baruzzi, W.G. Habashi, M. Fortin, J. Dompierre and M-G. Vallet, *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part II. Structured grids*, Int. J. Numer. Meth. Fluids, 39 (2002), pp. 657–673.

[2] L. Ambrosio and V.M. Tortorelli, *Approximation of functionals depending on jumps by elliptic functionals via* Γ*-convergence*, Comm. Pure Appl. Math., 43 (1990), pp. 999–1036.

[3] L. Ambrosio and V.M Tortorelli, *On the approximation of free discontinuity problems*, Boll. Un. Mat. Ital. B (7), 6 (1992), pp. 105–123.

[4] M. Artina, M. Fornasier, S. Micheletti and S. Perotto, *Anisotropic mesh adaptation for crack detection in brittle materials*, SIAM J. Sci. Comput., 37 (2015), pp. B633–B659.

[5] M. Artina, M. Fornasier, S. Micheletti and S. Perotto, *Anisotropic adaptive meshes for brittle fractures: parameter sensitivity*, In: Numerical Mathematics and Advanced Applications. Series: Lect. Notes Comput. Sci. Eng., Vol. 103, Springer, A. Abdulle, S. Deparis, D. Kressner, F. Nobile and M. Picasso (Eds.) (2015), pp. 293–302.

[6] M. Artina, M. Fornasier, S. Micheletti and S. Perotto, *The benefits of anisotropic mesh adaptation for brittle fractures under plane-strain conditions*, In: New Challenges in Grid Generation and Adaptivity for Scientific Computing, SEMA SIMAI Springer, Vol. 5, Springer Cham, S. Perotto and L. Formaggia (Eds.) (2015), pp. 43–67.

[7] Handbook of Medical Imaging, I.N. Bankman, Editor, Academic Press, Inc., Orlando, 2000.

[8] L. Biagiotti and C. Melchiorri, Trajectory Planning for Automatic Machines and Robots, Springer Publishing Company, Inc., 2008.

[9] B. Bourdin, *Image segmentation with a finite elements method*, M2AN Math. Model Numer. Anal., 33 (1999), pp. 229–244.

[10] S. Burke, C. Ortner and E. Süli, *An adaptive finite element approximation of a variational model of brittle fracture*, SIAM J. Numer. Anal., 48(3) (2010), pp. 980–1012.

[11] V. Caselles, R. Kimmel and G. Sapiro, *Geodesic active contours*, International Journal of Computer Vision. 22(1) (1997), pp.61–79.

[12] A. Chambolle, *Image segmentation by variational methods: Mumford and Shah functionals and the discrete approximations*, SIAM J. Appl. Math., 55(3) (1995), pp. 827–863.

[13] Ph. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[14] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numér., 2 (1975), pp. 77–84.

[15] G. Cortesani, *A finite element approximation of an image segmentation problem*, Math. Models Methods Appl. Sci., 9(2) (1999), pp.‘243–259.

[16] P. D'Ambra and G. Tartaglione, *Solution of Ambrosio–Tortorelli model for image segmentation by generalized relaxation method*, Commun. Nonlinear Sci. Numer. Simulat., 20 (2015), pp. 819–831.

[17] I. Dowman, K. Jacobsen, G. Konecny and R. Sandau High Resolution Optical Satellite Imagery, Whittles Publishing, 2012.

[18] L. Dedè, S. Micheletti and S. Perotto, *Anisotropic error control for environmental applications*, Appl. Numer. Math., 58 (2008), pp. 1320–1339.

[19] J. Dompierre, M-G. Vallet, Y. Bourgault, M. Fortin and W.G. Habashi, *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD solutions. Part III: Unstructured grids*, Int. J. Numer. Meth. Fluids, 39 (2002), pp. 675–702.

[20] Z.L. Feng, J.W. Yin, G. Chen and J.X. Dong, *A finite element segmentation algorithm for Jacquard image*, J. Softw., 16(1) (2005), pp.‘58–66.

[21] N. Ferro, S. Micheletti and S. Perotto, *Anisotropic mesh adaptation for crack propagation induced by a thermal shock in 2D*, Comput. Methods Appl. Mech. Engrg., 331 (2018), pp. 138–158.

[22] J.M. Fried, *Multichannel image segmentation using adaptive finite elements*, Comput. Vis. Sci., 12(3) (2009); pp.‘125–135.

[23] L. Formaggia and S. Perotto, *New anisotropic a priori error estimates*, Numer. Math., 89(4) (2001), pp. 641–667.

[24] L. Formaggia and S. Perotto, *Anisotropic error estimates for elliptic problems*, Numer. Math., 94 (2003), pp. 67–92.

[25] W.G. Habashi, M. Fortin, M-G. Vallet, J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia and A. Tam, *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD solutions. Part I: Theory*, Int. J. Numer. Meth. Fluids, 32 (2000), pp. 725–744.

[26] F. Hecht, *New developement in Freefem++*, J. Numer. Math., 20(3-4) (2012), pp. 251–265.

[27] J.B. MacQueen, *Some methods for classification and analysis of multivariate observations*, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, University of California Press., (1967), pp. 281–297.

[28] D. Mery, Computer Vision for X-Ray Testing: Imaging, Systems, Image Databases, and Algorithms, Springer Publishing Company, Inc., 2015.

[29] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.

[30] S. Micheletti, S. Perotto and M. Picasso, *Stabilized finite elements on anisotropic meshes: a priori error estimates for the advection-diffusion and the Stokes problems*, SIAM J. Numer. Anal., 41(3) (2003), pp. 1131–1162.

[31] S. Micheletti and S. Perotto, *Output functional control for nonlinear equations driven by anisotropic mesh adaption: The Navier–Stokes equations*, SIAM J. Sci. Comput., 30 (2008), pp. 2817–2854.

[32] S. Micheletti, S. Perotto and M. Signorini, *Anisotropic mesh adaptation for the generalized Ambrosio-Tortorelli functional with application to brittle fracture*, Comput. Math. Appl., 75(6) (2018), pp. 2134–2152.

[33] N. Otsu, *A threshold selection method from gray-level histograms*, IEEE Trans. Sys., Man., Cyber., 9(1) (1979), pp. 62–66.

[34] N. Papucci, *Adattazione anisotropa di griglia applicata alla segmentazione di immagini*, Master Thesis in Mathematical Engineering, Politecnico di Milano, a.y. 2010/2011.

[35] M. Petrou and P. Bosdogianni, Image Processing the Fundamentals, Wiley, UK, 2004.

[36] R. Rodríguez, *Some remarks on Zienkiewicz-Zhu estimator*, Numer. Methods Partial Differential Equations, 10(5) (1994), pp. 625–635.

[37] J.A. Sethian, Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Materials Science, Cambridge University Press, 1999.

[38] M. Tuceryan and A.K. Jain, Handbook of Pattern Recognition & Computer Vision, C.H. Chen, L.F. Pau and P.S.P. Wang, Editors, World Scientific Publishing Co., Inc., River Edge, 1993.

[39] S.E. Umbaugh, Digital Image Processing and Analysis: Human and Computer Vision Applications with CVIPtools, 2nd ed., Boca Raton, CRC Press, 2010.

[40] M. Van Dyke, An Album of Fluid Motion, The Parabolic Press, Stanford, 1988.

[41] F.S. Yaacobson and D. Givoli, *An adaptive finite element procedure for the image segmentation problem* Comm. Numer. Methods Engrg., 14(7) (1998), pp. 621–632.

[42] O.C. Zienkiewicz and J.Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Int. J. Numer. Methods Engrg., 24(2) (1987), pp. 337–357.

[43] O.C. Zienkiewicz and J.Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique*, Int. J. Numer. Methods Engng, 33 (1992), pp. 1331–1364.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

38/2018 Domanin, M.; Gallo, D.; Vergara, C.; Biondetti, P.; Forzenigo, L.V.; Morbiducci, U.
Prediction of long term restenosis risk after surgery in the carotid bifurcation by hemodynamic and geometric analysis

39/2018 Ferro, N.; Micheletti, S.; Perotto, S.
Density-based inverse homogenization with anisotropically adapted elements

37/2018 Bonaventura, L.; Della Rocca A.;
Convergence analysis of a cell centered finite volume diffusion operator on non-orthogonal polyhedral meshes

32/2018 Dal Santo, N.; Deparis, S.; Manzoni, A.; Quarteroni, A.
An algebraic least squares reduced basis method for the solution of nonaffinely parametrized Stokes equations

34/2018 Laurino, F.; Coclite, A.; Tiozzo, A.; Decuzzi, P.; Zunino, P.;
A muliscale computational approach for the interaction of functionalized nanoparticles with the miscrovasculature

35/2018 Possenti, L.; Casagrande, G.; Di Gregorio, S.; Zunino, P.; Costantino, M.L.
Numerical simulations of the microvascular fluid balance with a non-linear model of the lymphatic system

36/2018 Agosti, A.; Ambrosi, D.; Turzi, S.
Strain energy storage and dissipation rate in active cell mechanics

31/2018 Quarteroni, A.
The role of statistics in the era of big data: A computational scientist' perspective

30/2018 Ieva, F.; Palma, F.; Romo, J.
Bootstrap-based Inference for Dependence in Multivariate Functional Data

29/2018 Manzoni, A; Bonomi, D.; Quarteroni, A.
Reduced order modeling for cardiac electrophysiology and mechanics: new methodologies, challenges & perspectives