

# Augmented Lagrangian method for an Euler's elastica based segmentation model that promotes convex contours

Egil Bae <sup>\*</sup>, Xue-Cheng Tai <sup>†</sup> and Wei Zhu <sup>‡</sup>

## Abstract

*In this paper, we propose an image segmentation model where an  $L^1$  variant of the Euler's elastica energy is used as boundary regularization. An interesting feature of this model lies in its preference for convex segmentation contours. However, due to the high order and non-differentiability of Euler's elastica energy, it is nontrivial to minimize the associated functional. As in recent work on the ordinary  $L^2$ -Euler's elastica model in imaging, we propose using an augmented Lagrangian method to tackle the minimization problem. Specifically, we design a novel augmented Lagrangian functional that deals with the mean curvature term differently than in previous works. The new treatment reduces the number of Lagrange multipliers employed, and more importantly, it helps represent the curvature more effectively and faithfully. Numerical experiments validate the efficiency of the proposed augmented Lagrangian method and also demonstrate new features of this particular segmentation model, such as shape driven and data driven properties.*

## 1 Introduction

Image segmentation is one of the fundamental topics in the fields of image processing and computer vision. The aim is to partition an image region into sub-regions in order to capture meaningful objects. In the literature, numerous approaches have been proposed for this problem. For instance, in the classical snake and active contour model by Kass, Witkin, and Terzopoulos [22], segmentation contours are driven to object boundaries by internal forces, such as regularity, and external ones, like the intensity gradient. They proposed minimizing the following functional:

$$\mathbf{E}(\mathcal{C}) = \alpha \int_0^1 |\mathcal{C}'(s)|^2 ds + \beta \int_0^1 |\mathcal{C}''(s)| ds - \int_0^1 |\nabla f(\mathcal{C}(s))|^2 ds, \quad (1)$$

where  $f : \Omega \rightarrow \mathbb{R}^2$  represents a given image,  $\mathcal{C}(s) : [0, 1] \rightarrow \mathbb{R}^2$  is a parameterized curve, and  $\alpha, \beta$  are positive tuning parameters. The first two terms impose regularity assumptions on the contour while the third one imposes restrictions induced by the given image. In

---

<sup>\*</sup>Norwegian Defence Research Establishment, P.O. Box 25, NO-2027 Kjeller, Norway E-mail: Egil.Bae@ffi.no

<sup>†</sup>Department of Mathematics, University of Bergen, P.O. Box 7803, NO-5020 Bergen, Norway. Email: tai@math.uib.no.

<sup>‡</sup>Department of Mathematics, University of Alabama, Box 870350, Tuscaloosa, AL 35487. E-mail: wzhu7@bama.ua.edu.

fact, as  $f$  takes on large gradients along object boundaries, the functional  $E(\mathcal{C})$  will take on a small value when the active contour  $\mathcal{C}$  resides on these boundaries. Later on, in [7], Caselles, Kimmel, and Sapiro developed a variational segmentation model called geodesic active contours, where one needs to minimize the functional:

$$\mathbf{E}(\mathcal{C}) = \int_0^1 g_o(|\nabla f(\mathcal{C}(s))|)|\mathcal{C}'(s)|ds, \quad (2)$$

over all  $\mathcal{C}(s) \in S = \{\mathcal{C} : [0, 1] \rightarrow \mathbb{R}^2 : \mathcal{C}(s) \text{ is piecewise } C^1, \text{ and } \mathcal{C}(0) = \mathcal{C}(1)\}$ . The function  $g_o : [0, +\infty) \rightarrow [0, +\infty)$  is an edge detection function satisfying  $g_o(0) = 1$ ,  $\lim_{r \rightarrow +\infty} g_o(r) = 0$ , and which is also strictly decreasing. The edge detection function takes on values close to 1 in regions with homogenous grey intensity while being close to zero at boundaries with sharp intensity gradients. Therefore, the above functional attains its minimum value when the active contour  $\mathcal{C}(s)$  lies along edges in the image.

Another classical segmentation model was proposed by Mumford and Shah [30]. In this approach, the segmentation problem amounts to finding a piecewise smooth function which approximates the given image, where also the discontinuity set is restricted to be smooth. They proposed minimizing the following functional:

$$\mathbf{E}(u, \Gamma) = \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \lambda \int_{\Omega} (u - f)^2 dx + \mu \text{Length}(\Gamma) \quad (3)$$

with respect to both the clean image function  $u$  defined on  $\Omega$  and the discontinuity set  $\Gamma \subset \Omega$ .  $\lambda, \mu > 0$  are tuning parameters. The methodology has brought forth plenty of variational models in many topics of image processing, including segmentation, denoising, inpainting, etc. An interesting offspring of Mumford-Shah's model is Chan-Vese' model [8], where the discontinuity set  $\Gamma$  is restricted to be closed curves, leading to a segmentation of the image into two or several regions. Moreover, a nice feature of Chan-Vese's model lies in its treatment of the boundary  $\Gamma$  via level set functions [32] that can handle curves with topological changes easily.

The above mentioned segmentation models are all grey intensity based models, that is, the segmentation result mainly relies on grey intensity values of the images. However, due to the complexity of real images, meaningful objects might be occluded by other ones, or some parts of them may not be distinguished from the background by the contrast. For instance, in medical image analysis, target organs may be blended with other ones, and some parts of them may be occluded by other organs or even missing due to the imaging conditions. Therefore, those grey intensity based segmentation models might not be well suited for segmenting the target objects from such images. To overcome this issue, one possible way is to incorporate prior knowledge about the shapes of the target objects into the segmentation process. A lot of work have also focused on this research topic – segmentation using prior shape knowledge, see e.g. [12, 13, 14, 23]. Another possible way is to impose constraints on the segmentation contours to restore the boundaries that are missing or not well defined by the grey intensity gradient of the images. Of course, those well-known models discussed above all contain regularity terms on the segmentation contours. However, the main purpose of these terms is to improve the smoothness of the contours. They add no other geometric preference on the contours intentionally. In this work, we consider a new variant of Chan-Vese's model by using the  $L^1$ -Euler's elastica energy as regularization on the segmentation contours. As will be discussed later, the most interesting feature of this new

model lies in its capability of producing segmentations with convex contours. Moreover, the new segmentation model also possesses other new features – *data driven and shape driven properties*, the later of which has never been discussed in the literature, to the best of our knowledge.

Before introducing the  $L^1$ -Euler’s elastica energy, recall first the standard Euler’s elastica energy of the curve  $\Gamma$ :

$$\mathbf{E}(\Gamma) = \int_{\Gamma} (a + b\kappa^2) ds, \quad (4)$$

where  $\kappa$  represents the curvature of the curve and  $a, b > 0$  are two parameters. In the field of mathematical imaging, Euler’s elastica was first introduced by Nitzberg, Mumford, and Shiota in their famous work on segmentation with depth [31]. Since then, it has been adapted to many fundamental problems in imaging, such as image inpainting [11, 26], image restoration [1, 2, 39], and image segmentation [42, 45]. Notice that this elasticity energy contains a quadratic dependence on curvature. In fact, in [31] the authors considered another variant of Euler’s elastica that linearly depends on the magnitude of curvature, that is,  $\int_{\Gamma} (a + b|\kappa|) ds$ , based on the requirement of preserving corners in the segmentation contours. For the brevity of presentation, these two Euler’s elastica terms will be called the  $L^2$ - and  $L^1$ -Euler’s elastica terms respectively in the later sections.

Recently, a lot of research have focused on the development of fast and reliable numerical methods for minimizing curvature based functionals, such as the multigrid algorithm [5], the homotopy method [41], augmented Lagrangian method (ALM) based algorithms [39, 40, 45] and graph cut based algorithms [4, 18]. Another interesting line of research is the design of convex relaxation approaches [6, 29, 37, 38]. Their main advantage is the ability to avoid getting stuck in local minimums, but they also have several limitations: The computational cost and memory requirements are much higher due to the higher dimensional formulations of the problems; The curve must be represented discretely by a finite number of line segments; The convex relaxations only correspond approximately to the original problems, and in particular the tightest convex relaxation for the Euler’s elastica model without additional boundary length regularization reduces to constantly zero [6].

In this work, inspired by our previous work [45], we focus on the exact non-convex formulation and intend to construct a novel ALM based algorithm with fewer Lagrange multipliers, which markedly reduces the effort of choosing appropriate parameters for obtaining minimizers of the proposed model. Moreover, by employing fewer auxiliary variables, more direct connections among those variables are ensured, which leads to a more accurate discrete representation of the curvature of the segmentation contours. The augmented Lagrangian method was first used to get fast numerical schemes for nonlinear image processing models in [40], including the total variation based Rudin-Osher-Fatemi (ROF) model. Extensions of the augmented Lagrangian method for the ROF model to a spatially continuous setting were proposed in [21], and an interesting discussion about convergence properties were given. The fast algorithms were extended to  $L^2$ -Euler’s elastica and surface curvature models in [39, 44, 45]. For these non-linear higher order problems, there are many auxiliary variables to compute and constraints for the Lagrangian methods to handle. One of the contributions of this work is to reduce the number of auxiliary variables and constraints. We note that Myllykoski et al. [28] recently also considered an approach to reduce the number of Lagrange multipliers for an ALM based method for the surface curvature based image denoising model in [43].

As a summary, the new contribution of the current work lies in two aspects:

- We propose a curvature based segmentation model that promotes segmentations with convex contours;
- We develop an augmented Lagrangian algorithm for the new segmentation model, where curvature is treated in a novel and more accurate way, using fewer variables and parameters than in related work.

The rest of this paper is organized as follows. In the next section, we discuss the new segmentation model, where the  $L^1$ -Euler's elastica is used as regularization of the segmentation contours, and also present the specific application to segmentation of objects with convex contours. In Section 3, we discuss a natural extension of our previous work [45] for this new model and address several problematic issues, such as choosing appropriate parameters. After that, a novel ALM algorithm is proposed to ameliorate those issues. We then discuss its numerical implementation in Section 4, which is followed by numerical experiments that validate the features of the new model and the efficiency of the proposed ALM algorithm in Section 5. Section 6 is devoted to conclusions.

## 2 $L^1$ -Euler's elastica extension of Chan-Vese's model

In [45], we considered an extension of Chan-Vese's segmentation model by using the  $L^2$ -Euler's elastica as a new regularizer on the segmentation contours instead of the standard curve length regularization. This simple modification introduces new features of the resulting model: 1) the model is able to connect the boundary around missing parts of objects; 2) it can also capture objects of relatively large size while omitting those of small sizes; 3) it is more suitable than the original Chan-Vese's model for keeping elongated structures. These features have particular applications in medical imaging. For instance, MRI images often contain the mixture of organs or tissues of different sizes. By applying this model, it is possible to locate the boundaries of organs of considerable sizes while ignoring tiny tissues.

In this work, we intend to study another modification of Chan-Vese's segmentation model, that is, instead of the  $L^2$ -Euler's elastica, we utilize the  $L^1$ -Euler's elastica as regularization on the segmentation contour. With the aid of a level set function [32], we may express this new variational model (called *ECV- $L^1$*  segmentation model) as follows:

$$\mathbf{E}(\phi, c_1, c_2) = \int_{\Omega} (f - c_1)^2 H(\phi) + (f - c_2)^2 (1 - H(\phi)) + \int_{\Omega} \left[ a + b \left| \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} \right| \right] |\nabla H(\phi)|, \quad (5)$$

where  $f : \Omega \rightarrow \mathbb{R}$  is a given image,  $\phi$  is a level set function whose zero level set locates object boundaries, and  $a, b > 0$  are parameters. The term  $\nabla \cdot (\nabla \phi / |\nabla \phi|)$  denotes the mean curvature of the level curves of  $\phi$ , and the second part of the above functional just represents the  $L^1$ -Euler's elastica of the zero level set of  $\phi$ . The most remarkable feature of this new modified segmentation model is that it promotes convex contours provided the parameter  $b$  is sufficiently large. It may therefore be used for the particular task of segmenting objects that are known to have a convex shape. This specific feature is supported by the following theorem in differential geometry [16]:

**Theorem 1.** *Let  $\Gamma$  be any closed piecewise smooth curve in  $\mathbb{R}^2$  and  $\kappa$  its curvature, then  $\int_{\Gamma} |\kappa| ds \geq 2\pi$  and equality holds when  $\Gamma$  is a convex curve.*

We remind that a convex curve is a closed curve that is the boundary of a convex set. As this integral along convex curves is always  $2\pi$ , one still needs the length term to prohibit excessively large contours in the segmentation.

A nice feature of using a level set function in the above model is that since multiple curves can be represented by a single level set function  $\phi$ , one may impose convex shape priors on several objects simultaneously. Moreover, if the parameter  $b$  is chosen large, only those convex contours that encompass relatively large regions will be preserved in the final segmentation.

In contrast to the *ECV-L<sup>2</sup>* segmentation model introduced in [45], the *ECV-L<sup>1</sup>* segmentation model discussed in this paper assumes two new features: first, it prefers convex boundaries when the parameter  $b$  is chosen large; second, the model is capable of preserving sharp corners, since the  $L^1$ -Euler's elastica linearly depends on the magnitude of curvature, while the *ECV-L<sup>2</sup>* segmentation model necessarily smears out corners. The corner preservation property of the  $L^1$ -Euler's elastica energy was first pointed out in [31].

Just as in many other higher order variational models [1, 2, 31, 43], the numerical treatment of this new model is also nontrivial. The difficulty is mainly due to the presence of the curvature term, which makes the model high order, nonlinear, and non-differentiable. Recently, quite a few research works have focused on the development of efficient algorithms for these models, such as the multigrid algorithms [5], the homotopy method [41], and augmented Lagrangian methods (ALMs) [40, 39, 45]. In this work, we confine ourself to ALM in order to develop an efficient and reliable fast algorithm for the functional (Eq. 5), which will be detailed in the following sections.

### 3 The development of a novel augmented Lagrangian method

In this section, we first consider a natural extension of our previous work [45] for the new segmentation model and point out its problematic issues, and then develop a novel ALM for the minimization of the original model (Eq. 5).

#### 3.1 Issues arising from the previously proposed ALM for the model

Recently, ALMs have proven to be very useful in designing fast algorithms for variational imaging models with higher order terms or/and non-differential terms [39, 40, 44]. These models include the classical Rudin-Osher-Fatemi (ROF) model [36] and the Euler's elastica based models [2, 11]. The idea of ALMs is first to convert the minimization of those functionals to constrained problems, from which appropriate augmented Lagrangian functionals can be constructed. Based on the theory of optimization, the search of minimizers of the original functional amounts to finding saddle points of the resulting Lagrangian functionals. To get saddle points, one only needs to deal with several lower order subproblems, some of which can be solved using FFT while others have closed-form solutions. Therefore, with the utilization of ALMs, those issues relating to higher order terms and non-differentiable terms can be solved or avoided.

Following the idea in the work [10], we may introduce a new function  $u = H(\phi)$ , and as  $\nabla \cdot (\nabla H(\phi)/|\nabla H(\phi)|) = \nabla \cdot (\nabla u/|\nabla u|)$ , one can rewrite the functional (5) as follows:

$$\mathbf{E}(u, c_1, c_2) = \int_{\Omega} (f - c_1)^2 u + (f - c_2)^2 (1 - u) + \int_{\Omega} \left[ a + b \left| \nabla \cdot \frac{\nabla u}{|\nabla u|} \right| \right] |\nabla u|, \quad (6)$$

where  $u$  is supposed to take on the value of either 0 or 1. Note also that the curvature makes pointwise sense only for  $C^2$  functions. By following the ideas of [10, 45], one could relax the value of  $u$  inside the interval  $[0, 1]$ , introduce some auxiliary variables, and thus get the new constrained minimization problem:

$$\begin{aligned} \min_{u, \mathbf{p}, \mathbf{n}, c_1, c_2} \int_{\Omega} [(f - c_1)^2 u + (f - c_2)^2 (1 - u)] + \int_{\Omega} [a + b|\nabla \cdot \mathbf{n}|] |\mathbf{p}|, \\ \text{with } \mathbf{p} = \nabla u, \mathbf{n} = \nabla u / |\nabla u|, u \in [0, 1]. \end{aligned} \quad (7)$$

To deal with this minimization problem, just as in our previous work [45], one could construct an augmented Lagrangian functional as follows:

$$\begin{aligned} \mathcal{L}(v, u, q, \mathbf{p}, \mathbf{n}, \mathbf{m}, c_1, c_2; \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) &= \int_{\Omega} [(f - c_1)^2 v + (f - c_2)^2 (1 - v)] \\ &+ \int_{\Omega} (a + b|q|) |\mathbf{p}| \\ &+ r_1 \int_{\Omega} (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \int_{\Omega} \lambda_1 (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) \\ &+ \frac{r_2}{2} \int_{\Omega} |\mathbf{p} - \nabla u|^2 + \int_{\Omega} \lambda_2 \cdot (\mathbf{p} - \nabla u) \quad (8) \\ &+ \frac{r_3}{2} \int_{\Omega} (q - \nabla \cdot \mathbf{n})^2 + \int_{\Omega} \lambda_3 (q - \nabla \cdot \mathbf{n}) \\ &+ \frac{r_4}{2} \int_{\Omega} (v - u)^2 + \int_{\Omega} \lambda_4 (v - u) + \delta_{\mathcal{D}}(v) \\ &+ \frac{r_5}{2} \int_{\Omega} |\mathbf{n} - \mathbf{m}|^2 + \int_{\Omega} \lambda_5 \cdot (\mathbf{n} - \mathbf{m}) + \delta_{\mathcal{R}}(\mathbf{m}), \end{aligned}$$

where  $\mathcal{D} = [0, 1]$  and  $\mathcal{R} = \{\mathbf{m} \in L^2(\Omega) : |\mathbf{m}| \leq 1 \text{ a.e. in } \Omega\}$ , and  $\delta_{\mathcal{D}}(v)$  and  $\delta_{\mathcal{R}}(\cdot)$  are the characteristic functions on the sets  $\mathcal{D}$  and  $\mathcal{R}$  respectively:

$$\begin{aligned} \delta_{\mathcal{D}}(v) &= \begin{cases} 0, & \text{if } v \in \mathcal{D}; \\ +\infty, & \text{otherwise.} \end{cases} \\ \delta_{\mathcal{R}}(\mathbf{m}) &= \begin{cases} 0, & \text{if } \mathbf{m} \in \mathcal{R}; \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

The  $r'_i s, i = 1, \dots, 5$  are positive parameters, while  $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$  are Lagrange multipliers. As discussed in [39], a tricky technique used here is the introduction of the variable  $\mathbf{m}$ , which helps simplify the associated subproblem for the variable  $\mathbf{p}$ . In fact, as  $\mathbf{m} \in \mathcal{R}$ , that is,  $|\mathbf{m}| \leq 1$ , then  $|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m} \geq 0$  and the equality holds if and only if  $\mathbf{m} = \mathbf{p}/|\mathbf{p}|$ . One thus avoids the quadratic term  $\int_{\Omega} (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m})^2$ , which leads to a simpler subproblem for  $\mathbf{p}$ . Moreover, the introduction of the variable  $q$  is to tackle the non-differentiable term of mean curvature.

Using the same technique as in [45], the saddle points of this Lagrangian functional can be found by solving the resulting subproblems using either Fast Fourier Transform (FFT) or their closed-form solutions. However, to do so, one needs to choose appropriate auxiliary parameters  $r'_i s, i = 1, \dots, 5$ . In fact, the choice of these parameters for the current segmentation model is more involved than for denoising problems. This is because for denoising problems, the cleaned images should be close to the given noisy ones, which is an intrinsic constraint that is not possessed by segmentation problems.

Basically, the principle of choosing those auxiliary parameters naturally meets two fundamental requirements. First, segmentation results should be mainly determined by the grey intensity of the image under consideration, which give the regularization term a supportive role. Second, the quantity of mean curvature of segmentation contours needs to be expressed reliably, otherwise, the above functional (8) fails to faithfully reflect the essence of the  $ECV-L^1$  segmentation model (Eq. 5). To satisfy the first prerequisite, besides the model parameters  $a$  and  $b$ , those auxiliary parameters  $r'_i$ s need to be chosen relatively small so that the fitting terms remain dominant. On the other hand, to fulfill the second need, one has to set  $r'_i$ s relatively large in order to strengthen the connection among those auxiliary variables and thus express the curvature term effectively. These two seemingly opposite prerequisites considerably narrow the feasible range of the parameters  $r'_i$ s. Moreover, as discussed above, theoretically, the  $ECV-L^1$  segmentation model promotes convex contours, which also requires the curvature to be expressed accurately so that this particular feature of the model can be embodied. Therefore, the choice of these auxiliary parameters becomes a crucial issue for the numerical implementation of the  $ECV-L^1$  segmentation model.

To overcome this issue, one possible way is to utilize other methods to represent the curvature term directly, such as the well-known  $\Gamma$ -convergence [15]. In the literature, the notion of  $\Gamma$ -convergence has been successfully employed to deal with the minimization of functionals involving the  $L^2$ -Euler's elastica for image segmentation [20, 25]. With this technique, the original functional can be approximated by a sequence of more tractable functionals whose minimizers converge to that of the original functional. However, to the best of our knowledge, there is no related work on functionals involving the  $L^1$ -Euler's elastica.

In this paper, we still stick to augmented Lagrangian methods, but intend to develop a different augmented Lagrangian functional with fewer Lagrange multipliers for the  $ECV-L^1$  segmentation model. Besides considerably ameliorating the issue of choosing auxiliary parameters, due to the fewer new variables involved, this novel ALM can capture the mean curvature more directly and reliably.

### 3.2 A novel augmented Lagrangian method for the $ECV-L^1$ model

In this work, we adhere to the original form of the model (Eq. 5), that is, we will use a level set function instead of a binary function as in (Eq. 6), to represent the segmentation contour. This is because the mean curvature of the contour can be expressed more accurately in terms of a level set function than a binary function after discretization.

Besides keeping the level set function, we plan to develop a new augmented Lagrangian functional for the  $ECV-L^1$  segmentation model with fewer Lagrange multipliers, but with a more tight connection among those auxiliary variables that describe curvature. In [39], as discussed above, in the Euler's elastica based image denoising model, the curvature term  $\nabla \cdot (\nabla u / |\nabla u|)$  was expressed as  $\nabla \cdot \mathbf{n}$  with  $\mathbf{n} = \mathbf{m}$ ,  $\mathbf{m} = \mathbf{p}/|\mathbf{p}|$ , and  $\mathbf{p} = \nabla u$ . The introduction of  $\mathbf{m}$  is to relax the direct connection of  $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$ . This new variable considerably simplifies the treatment of the associated subproblem for  $\mathbf{p}$ . Due to the intrinsic constraint that denoised images should be close to the noisy ones, these relaxations often work very well for denoising problems [39]. However, these relaxations become problematic when dealing with curvature terms in segmentation problems. As discussed above, the main reason lies in the fact that with too many relaxations, the curvature terms often cannot be expressed accurately.

To fix this issue, we first approximate  $H(\phi)$  and  $\nabla H(\phi)$  as in [8] as:

$$\begin{aligned} H(\phi) &= \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right), \\ \nabla H(\phi) &= \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + \phi^2} \nabla \phi, \end{aligned}$$

where  $\epsilon > 0$  is a small parameter. We can then develop a new augmented Lagrangian functional with a direct connection between  $\mathbf{p}$  and  $\mathbf{n}$  for the original model (Eq. 5) as follows:

$$\begin{aligned} \mathcal{L}(\phi, q, \mathbf{p}, \mathbf{n}, c_1, c_2; \boldsymbol{\lambda}_1, \lambda_2, \boldsymbol{\lambda}_3) &= \int_{\Omega} \left\{ (f - c_1)^2 \left( \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) \right. \\ &+ \left. (f - c_2)^2 \left( \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) \right\} + \int_{\Omega} (a + b|q|) \frac{\epsilon}{\pi(\epsilon^2 + \phi^2)} |\mathbf{p}| \\ &+ \frac{r_1}{2} \int_{\Omega} |\mathbf{p} - \nabla \phi|^2 + \int_{\Omega} \boldsymbol{\lambda}_1 \cdot (\mathbf{p} - \nabla \phi) \tag{9} \\ &+ \frac{r_2}{2} \int_{\Omega} (q - \nabla \cdot \mathbf{n})^2 + \int_{\Omega} \lambda_2 (q - \nabla \cdot \mathbf{n}) \\ &+ \frac{r_3}{2} \int_{\Omega} \|\mathbf{p}\mathbf{n} - \mathbf{p}\|^2 + \int_{\Omega} \boldsymbol{\lambda}_3 \cdot (\mathbf{p}\mathbf{n} - \mathbf{p}). \end{aligned}$$

In this new functional, the variables  $\mathbf{p}$  and  $\mathbf{n}$  are connected through  $|\mathbf{p}\mathbf{n} - \mathbf{p}| = 0$ . This functional involves only three Lagrange multipliers, instead of five as the augmented Lagrangian functional (8), which considerably reduces the effort of choosing appropriate auxiliary parameters for the numerical implementation.

With this new Lagrangian functional (9), as discussed in [39, 45], one just needs to find its saddle points in order to get local minimizers of the original functional (5). To do so, we employ the standard iterative strategy: we minimize the corresponding functional for each variable by fixing the other ones, and then update the Lagrange multipliers; this process will be repeated until all the variables are convergent, indicating that some saddle point has been obtained. In what follows, we discuss the subproblems associated with all the

variables. Specifically, the corresponding functionals can be expressed as follows:

$$\varepsilon_1(q) = \int_{\Omega} b|q||\mathbf{p}| + \frac{r_2}{2} \int_{\Omega} (q - \nabla \cdot \mathbf{n})^2 + \int_{\Omega} \lambda_2(q - \nabla \cdot \mathbf{n}), \quad (10)$$

$$\begin{aligned} \varepsilon_2(\phi) &= \int_{\Omega} [(f - c_1)^2 \left( \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) + (f - c_2)^2 \left( \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right)] + \frac{r_1}{2} \int_{\Omega} |\mathbf{p} - \nabla\phi|^2 \\ &+ \int_{\Omega} \boldsymbol{\lambda}_1 \cdot (\mathbf{p} - \nabla\phi) \end{aligned} \quad (11)$$

$$\begin{aligned} \varepsilon_3(\mathbf{p}) &= \int_{\Omega} (a + b|q|)|\mathbf{p}| + \frac{r_1}{2} \int_{\Omega} |\mathbf{p} - \nabla\phi|^2 + \int_{\Omega} \boldsymbol{\lambda}_1 \cdot (\mathbf{p} - \nabla\phi) + \frac{r_3}{2} \int_{\Omega} \|\mathbf{p}|\mathbf{n} - \mathbf{p}\|^2 \\ &+ \int_{\Omega} \lambda_3(|\mathbf{p}|\mathbf{n} - \mathbf{p}), \end{aligned} \quad (12)$$

$$\varepsilon_4(\mathbf{n}) = \frac{r_2}{2} \int_{\Omega} (q - \nabla \cdot \mathbf{n})^2 + \int_{\Omega} \lambda_2(q - \nabla \cdot \mathbf{n}) + \frac{r_3}{2} \int_{\Omega} \|\mathbf{p}|\mathbf{n} - \mathbf{p}\|^2 + \int_{\Omega} \lambda_3(|\mathbf{p}|\mathbf{n} - \mathbf{p}), \quad (13)$$

$$\varepsilon_5(c_1) = \int_{\Omega} (f - c_1)^2 \left( \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right), \quad (14)$$

$$\varepsilon_6(c_2) = \int_{\Omega} (f - c_2)^2 \left( \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right). \quad (15)$$

As discussed in [39, 45], the minimizers of the functionals  $\varepsilon_1(q)$ ,  $\varepsilon_5(c_1)$ , and  $\varepsilon_6(c_2)$  can be obtained explicitly, and the ones for  $\varepsilon_2(\phi)$  and  $\varepsilon_4(\mathbf{n})$  are determined by the corresponding Euler-Lagrange equations. For the functional  $\varepsilon_3(\mathbf{p})$ , it possesses a non-differential term involving  $|\mathbf{p}|\mathbf{p}$  and therefore its Euler-Lagrange equation can only be obtained for its regularized version, which introduces approximation errors and makes the Euler-Lagrange equation intractable numerically. This explains the motivation of introducing the new variable  $\mathbf{m}$  in [39]. In this work, we try to find its minimizer directly. Let's first reformulate the functional  $\varepsilon_3(\mathbf{p})$  as follows:

$$\begin{aligned} \varepsilon_3(\mathbf{p}) &= \int_{\Omega} \left[ (a + b|q|) \frac{\epsilon}{\pi(\epsilon^2 + \phi^2)} + \boldsymbol{\lambda}_3 \cdot \mathbf{n} \right] |\mathbf{p}| + \frac{r_1 + r_3(1 + |\mathbf{n}|^2)}{2} \left| \mathbf{p} - \frac{\boldsymbol{\lambda}_3 + r_1 \nabla\phi - \boldsymbol{\lambda}_1}{r_1 + r_3(1 + |\mathbf{n}|^2)} \right|^2 \\ &- \int_{\Omega} r_3 \mathbf{p} \cdot \mathbf{n} |\mathbf{p}| + \tilde{c}, \end{aligned} \quad (16)$$

where  $\tilde{c}$  is independent of  $\mathbf{p}$ . Note that there is no spatial derivative of  $\mathbf{p}$  in  $\varepsilon_3(\mathbf{p})$ . Therefore, one just needs to consider the minimizer of its integrand at each point of the region  $\Omega$ . Specifically, we only need to find the minimizer of the function  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  as follows:

$$g(\mathbf{x}) = \lambda|\mathbf{x}| + \frac{\mu}{2} |\mathbf{x} - \mathbf{a}|^2 + (\nu \cdot \mathbf{x})|\mathbf{x}|, \quad (17)$$

where  $\lambda, \mu \in \mathbb{R}$ ,  $\mathbf{a}, \nu \in \mathbb{R}^2$  are given with  $\mu > 0$ . To this end, we propose the following theorem.

**Theorem 2.** *Assume that  $\mu > 2|\nu|$ . Let  $\theta$  be the angle between the vector  $\mathbf{a}$  and the minimum vector of  $g(\mathbf{x})$  listed in (17), and  $\alpha$  the angle between  $\mathbf{a}$  and  $\nu$ . Then the following arguments hold:*

- if  $\lambda \geq \mu|\mathbf{a}|$ , then  $g(\mathbf{x})$  attains its minimum at  $\mathbf{x} = 0$ .

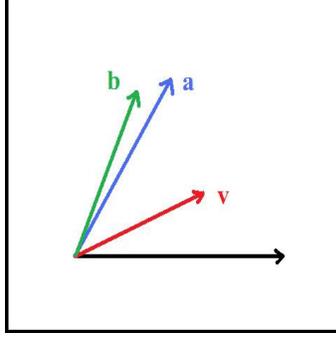


Figure 1: Two vectors  $\nu$ ,  $\mathbf{a}$  and a possible unit vector  $\mathbf{b}$  that maximizes the function  $\psi(\mathbf{b})$ .

- if  $\lambda < \mu|\mathbf{a}|$ , the minimum point of  $g(\mathbf{x})$  can be determined according to the following four cases:

1. if  $\mathbf{a} = \nu = 0$ , the minimum occurs at  $\mathbf{x} = 0$  if  $\lambda \geq 0$  and any vector of length  $-\frac{\lambda}{\mu}$  if  $\lambda < 0$ ;

2. if  $\mathbf{a} \neq 0, \nu = 0$ , the minimum point can be expressed as  $(1 - \frac{\lambda}{\mu|\mathbf{a}|})\mathbf{a}$ ;

3. if  $\mathbf{a} = 0, \nu \neq 0$ , the minimum occurs at  $\frac{\lambda}{\mu - 2|\nu|} \frac{\nu}{|\nu|}$ ;

4. if  $\mathbf{a} \neq 0, \nu \neq 0$ , the angles  $\theta$  and  $\alpha$  satisfy the equation

$$\mu^2|\mathbf{a}| \sin \theta - \mu|\nu||\mathbf{a}| \sin \theta \cos(\theta + \alpha) + \lambda|\nu| \sin(\theta + \alpha) - \mu|\mathbf{a}||\nu| \sin \alpha = 0, \quad (18)$$

and  $g(\mathbf{x})$  has its minimum at  $\frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]\mathbf{b}}{\mu + 2\nu \cdot \mathbf{b}}$  with  $\mathbf{b}$  being a unit vector satisfying

$$\mathbf{b} = \frac{1}{|\mathbf{a}|} \begin{bmatrix} \cos \tilde{\theta} & -\sin \tilde{\theta} \\ \sin \tilde{\theta} & \cos \tilde{\theta} \end{bmatrix} \mathbf{a} \quad (19)$$

and  $\tilde{\theta} = \theta$  if  $\det[\nu \ \mathbf{a}] \geq 0$ ,  $\tilde{\theta} = -\theta$  if  $\det[\nu \ \mathbf{a}] < 0$ , where  $[\nu \ \mathbf{a}]$  denotes the  $2 \times 2$  matrix with the vectors  $\nu$  and  $\mathbf{a}$  being the first and second column respectively.

*Proof.* Set  $\mathbf{x} = t\mathbf{b}$ , where  $t \geq 0$  and  $|\mathbf{b}| = 1$ , and fix the vector  $\mathbf{b}$ , then  $g(\mathbf{x})$  becomes a function of  $t$  as follows:

$$\begin{aligned} h(t) &= \lambda t + \frac{\mu}{2}[t^2 - 2t(\mathbf{b} \cdot \mathbf{a}) + |\mathbf{a}|^2] + t^2\nu \cdot \mathbf{b} \\ &= \left(\frac{\mu}{2} + \nu \cdot \mathbf{b}\right) \left[t - \frac{\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda}{\mu + 2\nu \cdot \mathbf{b}}\right]^2 + \frac{\mu}{2}|\mathbf{a}|^2 - \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]^2}{2(\mu + 2\nu \cdot \mathbf{b})} \end{aligned} \quad (20)$$

Since  $\mu > 2|\nu|$ ,  $\frac{\mu}{2} + \nu \cdot \mathbf{b} > 0$  for any unit vector  $\mathbf{b}$ . Note that  $h(t)$  is defined for  $t \geq 0$ . Therefore if  $t^* = \frac{\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda}{\mu + 2\nu \cdot \mathbf{b}} \leq 0$ ,  $h(t)$  attains its minimum value at  $t = 0$ ; while if  $t^* > 0$ ,  $h(t)$  takes on its minimum value  $\frac{\mu}{2}|\mathbf{a}|^2 - \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]^2}{2(\mu + 2\nu \cdot \mathbf{b})}$  at  $t^*$ .

Hence, if  $\lambda \geq \mu|\mathbf{a}|$ , then  $\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda \leq 0$  for any unit vector  $\mathbf{b}$ , and  $t^* \leq 0$ . This leads to our first claim in the theorem.

On the other hand, if  $\lambda < \mu|\mathbf{a}|$ , for any unit vector  $\mathbf{b}$ , one has either  $\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda \leq 0$  or  $\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda > 0$ . For the first case,  $t^* \leq 0$ , and  $h(t)$  takes its minimum value  $\frac{\mu}{2}|\mathbf{a}|^2$ , while for the second case, the minimum value reads  $\frac{\mu}{2}|\mathbf{a}|^2 - \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]^2}{2(\mu + 2\nu \cdot \mathbf{b})}$ , which is strictly less than  $\frac{\mu}{2}|\mathbf{a}|^2$ . Therefore, to find the minimum of  $g(\mathbf{x})$ , one needs to find the maximum of the function  $\psi(\mathbf{b}) = \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]^2}{2(\mu + 2\nu \cdot \mathbf{b})}$  over those unit vector  $\mathbf{b}$  with  $\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda > 0$ .

In what follows, let's consider those cases under the assumption  $\lambda < \mu|\mathbf{a}|$  one by one.

If  $\mathbf{a} = \nu = 0$ , then  $g(\mathbf{x}) = \lambda|\mathbf{x}| + \frac{\mu}{2}|\mathbf{x}|^2$ . As  $\mu > 0$  and  $\lambda < \mu|\mathbf{a}| = 0$ , simple calculation shows that the minimum value occurs at all those  $\mathbf{x}$  with length  $-\lambda/\mu$ .

If  $\mathbf{a} \neq 0, \nu = 0$ , then  $g(\mathbf{x})$  degenerates to be  $\lambda|\mathbf{x}| + \frac{\mu}{2}|\mathbf{x} - \mathbf{a}|^2$ . As shown in [39, 45], its minimum reads  $\max\{0, 1 - \frac{\lambda}{\mu|\mathbf{a}|}\}\mathbf{a}$ , and therefore  $(1 - \frac{\lambda}{\mu|\mathbf{a}|})\mathbf{a}$  since  $\lambda < \mu|\mathbf{a}|$ .

If  $\mathbf{a} = 0, \nu \neq 0$ ,  $\psi(\mathbf{b})$  degenerates to be  $\psi(\mathbf{b}) = \frac{\lambda^2}{2(\mu + 2\nu \cdot \mathbf{b})}$ , whose maximum is the unit vector  $-\frac{\nu}{|\nu|}$ . Therefore, the minimum of  $g(\mathbf{x})$  reads  $t^*(-\frac{\nu}{|\nu|}) = \frac{\lambda}{\mu - 2|\nu|} \frac{\nu}{|\nu|}$ .

If  $\mathbf{a} \neq 0, \nu \neq 0$ , as the angle between the vectors  $\mathbf{b}$  and  $\mathbf{a}$  is denoted by  $\theta$ , the function  $\psi(\mathbf{b})$  can be reformulated as a function  $\phi(\theta)$

$$\phi(\theta) = \frac{(\mu|\mathbf{a}| \cos \theta - \lambda)^2}{\mu + 2|\nu| \cos(\theta + \alpha)},$$

and simple calculation gives that  $\phi(\theta)$  takes its maximum value at the angle  $\theta$  satisfying the following equation

$$\mu^2|\mathbf{a}| \sin \theta - \mu|\nu||\mathbf{a}| \sin \theta \cos(\theta + \alpha) + \lambda|\nu| \sin(\theta + \alpha) - \mu|\mathbf{a}||\nu| \sin \alpha = 0. \quad (21)$$

Once the minimum  $\theta$  of  $\phi(\theta)$  is obtained, one still needs to determine which  $\mathbf{b}$  should be chosen as the minimum of  $\psi(\mathbf{b})$  since  $\theta$  is the angle between  $\mathbf{b}$  and  $\mathbf{a}$ . Notice that  $\psi(\mathbf{b}) = \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]^2}{2(\mu + 2\nu \cdot \mathbf{b})}$ . Geometrically, as shown in Fig. 1, the smaller the angle  $\theta$  is, the bigger the numerator of  $\psi(\mathbf{b})$  will be; while for its denominator, if the angle between  $\mathbf{b}$  and  $\nu$  is larger, the denominator will be smaller. These facts show that to attain the maximum value of  $\psi(\mathbf{b})$ ,  $\mathbf{b}$  should sit on the further side of  $\mathbf{a}$  with respect to  $\nu$ . This requires to use the sign of the determinant of the  $2 \times 2$  matrix  $[\nu \ \mathbf{a}]$  to determine the orientation of  $\mathbf{b}$  with respect to  $\mathbf{a}$ . Specifically, if  $\det[\nu \ \mathbf{a}] \geq 0$ , we may rotate the unit vector  $\mathbf{a}/|\mathbf{a}|$  counter-clockwisely to get  $\mathbf{b}$ ; while if  $\det[\nu \ \mathbf{a}] < 0$ ,  $\mathbf{b}$  can be obtained by rotating the vector clock-wisely. In this manner, one finds the maximum  $\mathbf{b}$  for  $\psi(\mathbf{b})$ , and thus gets the minimum of the goal function  $g(\mathbf{x})$  as  $t^*\mathbf{b} = \frac{[\mu(\mathbf{b} \cdot \mathbf{a}) - \lambda]}{\mu + 2\nu \cdot \mathbf{b}}\mathbf{b}$ , which proves the theorem.  $\square$

**Remark 1.** Notice that if one expresses the integrand of  $\varepsilon_3(\mathbf{p})$  as the function  $g(\mathbf{p})$  defined as (17), those parameters read  $\lambda = (a + b|q|)\epsilon/[\pi(\epsilon^2 + \phi^2)] + \boldsymbol{\lambda}_3 \cdot \mathbf{n}$ ,  $\mu = r_1 + r_3(1 + |\mathbf{n}|^2)$ , and  $\nu = -r_3\mathbf{n}$ . Then it is easy to see that the assumption  $\mu > 2|\nu|$  is surely satisfied, since  $r_3(1 + |\mathbf{n}|^2) \geq 2r_3|\mathbf{n}|$  for any  $\mathbf{n}$ . Therefore, the above theorem can be applied for the minimization of the functional  $\varepsilon_3(\mathbf{p})$ .

**Remark 2.** Based on this theorem, the closed-form minimizer of  $\varepsilon_3(\mathbf{p})$  only exist for several extreme cases. In general, one has to solve the equation (18). In this paper, we solve it using Newton's method with an initial  $\theta$  satisfying  $[\cos \theta, \sin \theta] = \mathbf{a}/|\mathbf{a}|$ , that is, choose the unit

vector along with the vector  $\mathbf{a}$  as the initial guess of  $\mathbf{b}$  when maximizing the function  $\psi(\mathbf{b})$ . In fact, if the minimum  $\theta$  is close to 0, by using the approximation that  $\cos(\theta + \alpha) \approx \cos \alpha$  and  $\sin(\theta + \alpha) \approx \sin \alpha$ , one gets an approximation

$$\sin \theta = \frac{|\nu|(\mu|\mathbf{a}| - \lambda) \sin \alpha}{\mu|\mathbf{a}|(\mu + |\nu| \cos \alpha)}, \quad (22)$$

and then we may also use this particular  $\theta$  as the initialization when solving (18) if  $\lambda \in [0, 2\mu|\mathbf{a}|]$ . This is because if  $\lambda \in [0, 2\mu|\mathbf{a}|]$  and also  $\mu > 2|\nu|$ , then  $|\mu|\mathbf{a}| - \lambda| \leq \mu|\mathbf{a}|$  and  $|\nu| < \mu + |\nu| \cos \alpha$ , and therefore one can guarantee that this approximation of  $\sin \theta$  lies inside  $[-1, 1]$ .

**Remark 3.** In fact, the constraint  $|\mathbf{p}|\mathbf{n} - \mathbf{p} = 0$  was also employed in [17]. However, in this work, we propose a novel method to deal with the associated subproblem of  $\mathbf{p}$  directly, while in [17], a fixed-point based technique was used. Specifically, during the iterative process, this constraint is relaxed to be  $|\mathbf{p}^{k-1}|\mathbf{n} - \mathbf{p} = 0$  in the  $k^{\text{th}}$  iteration, and the minimization of the subproblem of  $\mathbf{p}$  is circumvented and an approximation of its minimizer is obtained using the same technique as in [39].

Similarly as was done in [39, 45], as  $\varepsilon_1(q)$  can be rewritten as

$$\varepsilon_1(q) = \int_{\Omega} \frac{b\epsilon|\mathbf{p}|}{\pi(\epsilon^2 + \phi^2)} |q| + \frac{r_2}{2} \left[ q - \left( \nabla \cdot \mathbf{n} - \frac{\lambda_2}{r_2} \right) \right]^2 + \tilde{c}_1, \quad (23)$$

where  $\tilde{c}_1$  is independent of  $q$ , its minimizer reads

$$\text{Argmin}_q \varepsilon_1(q) = \max \left\{ 0, 1 - \frac{b\epsilon|\mathbf{p}|}{r_2\pi(\epsilon^2 + \phi^2)|q^*} \right\} q^* \quad (24)$$

with  $q^* = \nabla \cdot \mathbf{n} - \lambda_2/r_2$ . And the minimizers of  $\varepsilon_5(c_1)$ , and  $\varepsilon_6(c_2)$  are given as follows:

$$c_1 = \frac{\int_{\Omega} \left( \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) f}{\int_{\Omega} \left( \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right)}, \quad (25)$$

$$c_2 = \frac{\int_{\Omega} \left( \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right) f}{\int_{\Omega} \left( \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right)}. \quad (26)$$

As for the functionals  $\varepsilon_2(\phi)$  and  $\varepsilon_4(\mathbf{n})$ , standard procedures lead to the following Euler-Lagrange equations:

$$\begin{aligned} -r_1 \Delta \phi &= -\frac{\epsilon}{\pi(\epsilon^2 + \phi^2)} [(f - c_1)^2 - (f - c_2)^2] + (a + b|q|) \frac{2\epsilon|\mathbf{p}|\phi}{(\epsilon^2 + \phi^2)^2} \\ &\quad - \nabla \cdot (r_1 \mathbf{p} + \boldsymbol{\lambda}_1), \end{aligned} \quad (27)$$

$$-r_2 \nabla (\nabla \cdot \mathbf{n}) + r_3 |\mathbf{p}|^2 \mathbf{n} = -\nabla (r_2 q + \lambda_2) - (\lambda_3 - r_3 |\mathbf{p}|) \mathbf{p}. \quad (28)$$

In summary, as discussed above, to find saddle points of the new augmented Lagrangian functional (9), we minimize each of the above functionals  $\varepsilon'_i$ ,  $i = 1, \dots, 6$  by fixing other variables, and once all these variables are updated, we then advance the Lagrange multipliers  $\boldsymbol{\lambda}_1, \lambda_2, \boldsymbol{\lambda}_3$  as follows:

$$\boldsymbol{\lambda}_1^{\text{new}} = \boldsymbol{\lambda}_1^{\text{old}} + r_1 (\mathbf{p} - \nabla \phi), \quad (29)$$

$$\lambda_2^{\text{new}} = \lambda_2^{\text{old}} + r_2 (q - \nabla \cdot \mathbf{n}), \quad (30)$$

$$\boldsymbol{\lambda}_3^{\text{new}} = \boldsymbol{\lambda}_3^{\text{old}} + r_3 (|\mathbf{p}|\mathbf{n} - \mathbf{p}). \quad (31)$$

Table 1: Augmented Lagrangian method for the  $ECV-L^1$  segmentation model.

<p>1. Initialization: <math>\phi^0, q^0, \mathbf{p}^0, \mathbf{n}^0</math>, and <math>\lambda_1^0, \lambda_2^0, \lambda_3^0</math>.  For <math>k \geq 1</math>, do the following steps (Step 2 <math>\sim</math> 4):</p> <p>2. Compute an approximate minimizer <math>(\phi^k, q^k, \mathbf{p}^k, \mathbf{n}^k)</math> of the augmented Lagrangian functional with the fixed Lagrangian multiplier <math>\lambda_1^{k-1}, \lambda_2^{k-1}, \lambda_3^{k-1}</math>:</p> $(\phi^k, q^k, \mathbf{p}^k, \mathbf{n}^k) \approx \operatorname{argmin} \mathcal{L}(\phi, q, \mathbf{p}, \mathbf{n}, \lambda_1^{k-1}, \lambda_2^{k-1}, \lambda_3^{k-1}). \quad (32)$ <p>3. Update the Lagrangian multipliers</p> $\begin{aligned} \lambda_1^k &= \lambda_1^{k-1} + r_1(\mathbf{p}^k - \nabla \phi^k) \\ \lambda_2^k &= \lambda_2^{k-1} + r_2(q^k - \nabla \cdot \mathbf{n}^k) \\ \lambda_3^k &= \lambda_3^{k-1} + r_3( \mathbf{p}  \mathbf{n}^k - \mathbf{p}^k), \end{aligned}$ <p>4. Measure the relative residuals and stop the iteration if they are smaller than a threshold <math>\epsilon_r</math>.</p>
---

We repeat this process until all the variables are convergent. The above procedure can be summarized as in Table 1 and Table 2.

## 4 Numerical Implementation

In this section, we present the details of solving the equations (27) and (28) and update the variables  $q, \mathbf{p}$  and  $c_1, c_2$  as well as the Lagrange multiplier  $\lambda_1, \lambda_2, \lambda_3$  during each iteration when applying the iterative algorithm for finding saddle points of the Lagrangian functional (9). As the numerics are very similar to what discussed in [39, 45], only key points are included below.

Let  $\Omega = \{(i, j) | 1 \leq i \leq M, 1 \leq j \leq N\}$  be the discretized image domain and each point

Table 2: Alternating minimization method for solving the subproblems.

<p>1. Initialization: <math>\tilde{\phi}^0 = \phi^{k-1}, \tilde{q}^0 = q^{k-1}, \tilde{\mathbf{p}}^0 = \mathbf{p}^{k-1}</math>, and <math>\tilde{\mathbf{n}}^0 = \mathbf{n}^{k-1}</math>.</p> <p>2. For fixed Lagrangian multiplier <math>\lambda_1 = \lambda_1^{k-1}, \lambda_2 = \lambda_2^{k-1}</math>, and <math>\lambda_3 = \lambda_3^{k-1}</math>, solve the following subproblems :</p> $\tilde{\phi}^1 = \operatorname{argmin} \mathcal{L}(\phi, \tilde{q}^0, \tilde{\mathbf{p}}^0, \tilde{\mathbf{n}}^0, \lambda_1, \lambda_2, \lambda_3) \quad (33)$ $\tilde{q}^1 = \operatorname{argmin} \mathcal{L}(\tilde{\phi}^1, q, \tilde{\mathbf{p}}^0, \tilde{\mathbf{n}}^0, \lambda_1, \lambda_2, \lambda_3) \quad (34)$ $\tilde{\mathbf{p}}^1 = \operatorname{argmin} \mathcal{L}(\tilde{\phi}^1, \tilde{q}^1, \mathbf{p}, \tilde{\mathbf{n}}^0, \lambda_1, \lambda_2, \lambda_3) \quad (35)$ $\tilde{\mathbf{n}}^1 = \operatorname{argmin} \mathcal{L}(\tilde{\phi}^1, \tilde{q}^1, \tilde{\mathbf{p}}^1, \mathbf{n}, \lambda_1, \lambda_2, \lambda_3) \quad (36)$ <p>3. <math>(\phi^k, q^k, \mathbf{p}^k, \mathbf{n}^k) = (\tilde{\phi}^1, \tilde{q}^1, \tilde{\mathbf{p}}^1, \tilde{\mathbf{n}}^1)</math>.</p>
---

$(i, j)$  is called a pixel point. All the variable functions are defined on these pixel points. We first introduce the discrete backward and forward differential operators with periodic boundary condition as follows:

$$\begin{aligned}\partial_1^- \phi(i, j) &= \begin{cases} \phi(i, j) - \phi(i-1, j), & 1 < i \leq M; \\ \phi(1, j) - \phi(M, j), & i = 1. \end{cases} \\ \partial_1^+ \phi(i, j) &= \begin{cases} \phi(i+1, j) - \phi(i, j), & 1 \leq i < M-1; \\ \phi(1, j) - \phi(M, j), & i = M. \end{cases} \\ \partial_2^- \phi(i, j) &= \begin{cases} \phi(i, j) - \phi(i, j-1), & 1 < j \leq N; \\ \phi(i, 1) - \phi(i, N), & j = 1. \end{cases} \\ \partial_2^+ \phi(i, j) &= \begin{cases} \phi(i, j+1) - \phi(i, j), & 1 \leq j < N; \\ \phi(i, 1) - \phi(i, N), & j = N. \end{cases}\end{aligned}$$

and the central difference operators and the gradient operators are defined accordingly as

$$\begin{aligned}\partial_1^c \phi(i, j) &= (\partial_1^- \phi(i, j) + \partial_1^+ \phi(i, j))/2, \\ \partial_2^c \phi(i, j) &= (\partial_2^- \phi(i, j) + \partial_2^+ \phi(i, j))/2, \\ \nabla^\pm \phi(i, j) &= (\partial_1^\pm \phi(i, j), \partial_2^\pm \phi(i, j)).\end{aligned}$$

We first discuss how to solve the equation (27) using FFT. As this equation contains nonlinear terms of  $\phi$ , we employ the frozen coefficient technique to solve it, and consider the following equation

$$\begin{aligned}-r_1 \Delta \phi + \delta_\phi \phi &= \delta_\phi \phi - \frac{\epsilon}{\pi(\epsilon^2 + \phi^2)} [(f - c_1)^2 - (f - c_2)^2] + (a + b|q|) \frac{2\epsilon|\mathbf{p}|\phi}{(\epsilon^2 + \phi^2)^2} \\ &\quad - \nabla \cdot (r_1 \mathbf{p} + \boldsymbol{\lambda}_1),\end{aligned}\tag{37}$$

where  $\delta_\phi > 0$  is a small parameter, and its discretization as follows:

$$\begin{aligned}-r_1 \text{div}^- \nabla^+ \phi + \delta_\phi \phi &= \delta_\phi \phi - \frac{\epsilon}{\pi(\epsilon^2 + \phi^2)} [(f - c_1)^2 - (f - c_2)^2] + (a + b|q|) \frac{2\epsilon|\mathbf{p}|\phi}{(\epsilon^2 + \phi^2)^2} \\ &\quad - \partial_1^- (r_1 p_1 + \lambda_{11}) - \partial_2^- (r_1 p_2 + \lambda_{12}),\end{aligned}\tag{38}$$

where  $\mathbf{p} = \langle p_1, p_2 \rangle$  and  $\boldsymbol{\lambda}_1 = \langle \lambda_{11}, \lambda_{12} \rangle$ . Let's denote the right side by the function  $g$  and apply the discrete Fourier transform  $\mathcal{F}$  for both sides. Notice that

$$\begin{aligned}\mathcal{F} \partial_1^\pm \phi(i, j) &= \pm (e^{\pm \sqrt{-1} z_i^1} - 1) \mathcal{F} \phi(i, j), \\ \mathcal{F} \partial_2^\pm \phi(i, j) &= \pm (e^{\pm \sqrt{-1} z_j^2} - 1) \mathcal{F} \phi(i, j),\end{aligned}$$

where  $z_i^1 = 2\pi(i-1)/M, i = 1, \dots, M, z_j^2 = 2\pi(j-1)/N, j = 1, \dots, N$ . We get

$$(-2r_1(\cos z_i^1 + \cos z_j^2 - 2) + \delta_\phi) \mathcal{F} \phi(i, j) = \mathcal{F} g(i, j).\tag{39}$$

Once  $\mathcal{F} \phi$  is calculated,  $\phi$  can be obtained using the discrete inverse Fourier transform.

As for the equation (28), notice that the coefficient  $r_3|\mathbf{p}|^2$  varies over  $\Omega$  and also might not always be positive. We hence consider the following equation

$$-r_2 \nabla (\nabla \cdot \mathbf{n}) + D \mathbf{n} = (D - r_3|\mathbf{p}|^2) \mathbf{n} - \nabla (r_2 q + \lambda_2) - (\lambda_3 - r_3|\mathbf{p}|) \mathbf{p},\tag{40}$$

where  $D = \max_{x \in \Omega} (r_3 |\mathbf{p}|^2 + \delta_{\mathbf{n}})$  with  $\delta_{\mathbf{n}}$  being a small positive number. As  $\mathbf{n} = \langle n_1, n_2 \rangle$ ,  $\mathbf{p} = \langle p_1, p_2 \rangle$ , the discretization of (40) leads to a system,

$$\begin{aligned} -r_2(\partial_1^+ \partial_1^- n_1 + \partial_1^+ \partial_2^- n_2) + Dn_1 &= (D - r_3 |\mathbf{p}|^2) n_1 - \partial_1^+ (r_2 q + \lambda_2) - (\lambda_3 - r_3 |\mathbf{p}|) p_1, \\ -r_2(\partial_2^+ \partial_1^- n_1 + \partial_2^+ \partial_2^- n_2) + Dn_2 &= (D - r_3 |\mathbf{p}|^2) n_2 - \partial_2^+ (r_2 q + \lambda_2) - (\lambda_3 - r_3 |\mathbf{p}|) p_2. \end{aligned}$$

Denote the two right sides by the functions  $h_1, h_2$  respectively. One applies the discrete Fourier transform to both sides and gets the following  $2 \times 2$  system for  $(\mathcal{F}n_1, \mathcal{F}n_2)$  at each pixel point  $(i, j)$ :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathcal{F}n_1(i, j) \\ \mathcal{F}n_2(i, j) \end{pmatrix} = \begin{pmatrix} \mathcal{F}h_1(i, j) \\ \mathcal{F}h_2(i, j) \end{pmatrix} \quad (41)$$

with

$$\begin{aligned} a_{11} &= D - r_2(e^{\sqrt{-1}z_i^1} - 1)(1 - e^{-\sqrt{-1}z_i^1}), \\ a_{12} &= -r_2(e^{\sqrt{-1}z_i^1} - 1)(1 - e^{-\sqrt{-1}z_j^2}), \\ a_{21} &= -r_2(e^{\sqrt{-1}z_j^2} - 1)(1 - e^{-\sqrt{-1}z_i^1}), \\ a_{22} &= D - r_2(e^{\sqrt{-1}z_j^2} - 1)(1 - e^{-\sqrt{-1}z_j^2}), \end{aligned}$$

and

$$\begin{aligned} h_1(i, j) &= (D - r_3 |\mathbf{p}|^2)(i, j) n_1(i, j) - \partial_1^+ (r_2 q + \lambda_2)(i, j) - (\lambda_3 - r_3 |\mathbf{p}|)(i, j) p_1(i, j) \\ h_2(i, j) &= (D - r_3 |\mathbf{p}|^2)(i, j) n_2(i, j) - \partial_2^+ (r_2 q + \lambda_2)(i, j) - (\lambda_3 - r_3 |\mathbf{p}|)(i, j) p_2(i, j). \end{aligned}$$

Notice that the determinant of the above  $2 \times 2$  matrix equals to  $D^2 - 2Dr_2(\cos z_i^1 + \cos z_j^2 - 2) > 0$ . One can easily solve the above system to get  $\mathcal{F}n_1$  and  $\mathcal{F}n_2$  and then apply the discrete inverse Fourier transform to find the new updated  $\mathbf{n} = (n_1, n_2)$ .

As for the update of the variable  $\mathbf{p}$  for each iteration, in general, one needs to solve the equation (18). In this work, we employ the Newton's method with an initial  $\theta$  that was discussed in Remark 3.2, that is,  $\theta$  satisfies  $[\cos \theta, \sin \theta] = \mathbf{a}/|\mathbf{a}|$  with  $\mathbf{a}$  being the vector defined in the function  $g(\mathbf{x})$  (Eq. 17) or  $\theta$  is the angle defined in Eq. (22).

As for the variable  $q$ , we first calculate  $q^*$  and then get the updated  $q$  as follows:

$$q^*(i, j) = \partial_1^- n_1(i, j) + \partial_2^- n_2(i, j) - \frac{\lambda_2(i, j)}{r_2},$$

and

$$q(i, j) = \max \left\{ 0, 1 - \frac{b\epsilon |\mathbf{p}(i, j)|}{r_2 \pi (\epsilon^2 + \phi^2(i, j)) |q^*(i, j)|} \right\} q^*(i, j), \quad (42)$$

with  $|\mathbf{p}(i, j)| = \sqrt{p_1^2(i, j) + p_2^2(i, j)}$ .

The two scalar variables  $c_1$  and  $c_2$  can be easily updated according to Eqs. (25) and (26) respectively. Moreover, we update all the Lagrangian multipliers as follows:

$$\begin{aligned} \lambda_{11}^{new}(i, j) &= \lambda_{11}^{old}(i, j) + r_1(p_1(i, j) - \partial_1^+ \phi(i, j)), \\ \lambda_{12}^{new}(i, j) &= \lambda_{12}^{old}(i, j) + r_1(p_2(i, j) - \partial_2^+ \phi(i, j)), \\ \lambda_2^{new}(i, j) &= \lambda_2^{old}(i, j) + r_2(q(i, j) - \partial_1^- n_1(i, j) - \partial_2^- n_2(i, j)), \\ \lambda_{31}^{new}(i, j) &= \lambda_{31}^{old}(i, j) + r_3(|\mathbf{p}(i, j)| n_1(i, j) - p_1(i, j)), \\ \lambda_{32}^{new}(i, j) &= \lambda_{32}^{old}(i, j) + r_3(|\mathbf{p}(i, j)| n_2(i, j) - p_2(i, j)), \end{aligned}$$

where  $\boldsymbol{\lambda}_1 = \langle \lambda_{11}, \lambda_{12} \rangle$ ,  $\boldsymbol{\lambda}_3 = \langle \lambda_{31}, \lambda_{32} \rangle$ .

## 5 Numerical Experiments

In this section, numerical results are presented to validate the features of the Euler’s elastica based segmentation model by applying the proposed ALM. Besides being capable of capturing convex contours, the proposed  $ECV-L^1$  segmentation model also encompass other interesting features – data and shape driven properties. The *shape driven property* is a new feature that has never been discussed for conventional segmentation models in the literature to our knowledge.

For each experiment, to monitor whether the iterative process using the proposed ALM converges to a saddle point of the Lagrangian functional (9), as in [39, 45], we calculate the following relative residuals:

$$(R_1^k, R_2^k, R_3^k) = \frac{1}{|\Omega|} (|\tilde{R}_1^k|_{L^1}, |\tilde{R}_2^k|_{L^1}, |\tilde{R}_3^k|_{L^1}), \quad (43)$$

with

$$\begin{aligned} \tilde{R}_1^k &= \mathbf{p}^k - \nabla \phi^k, \\ \tilde{R}_2^k &= q^k - \nabla \cdot \mathbf{n}^k, \\ \tilde{R}_3^k &= |\mathbf{p}^k| \mathbf{n}^k - \mathbf{p}^k, \end{aligned}$$

where  $|\cdot|_{L^1}$  is the  $L^1$ -norm on  $\Omega$  and  $|\Omega|$  is the area of the domain. The relative residuals (43) can be used as the stopping criterion for the process, that is, given a small threshold  $\epsilon_r > 0$ , once  $R_i^k < \epsilon_r$  for  $i = 1, 2, 3$  and for some  $k$ , the iteration process will be terminated.

For the purpose of convergence checking, we also track the relative errors of the Lagrange multipliers:

$$(L_1^k, L_2^k, L_3^k) = \left( \frac{|\lambda_1^k - \lambda_1^{k-1}|_{L^1}}{|\lambda_1^{k-1}|_{L^1}}, \frac{|\lambda_2^k - \lambda_2^{k-1}|_{L^1}}{|\lambda_2^{k-1}|_{L^1}}, \frac{|\lambda_3^k - \lambda_3^{k-1}|_{L^1}}{|\lambda_3^{k-1}|_{L^1}} \right), \quad (44)$$

and the relative error of the iterate  $\phi^k$

$$\frac{|\phi^k - \phi^{k-1}|_{L^1}}{|\phi^{k-1}|_{L^1}}. \quad (45)$$

For the purpose of presentation, all the above quantities are shown in *log*-scale in the following results. Moreover, in all the experiments in this work, fixed parameters include  $\epsilon = 1.0$ ,  $\delta_u = \delta_n = 10^{-2}$ . And for the convenience of choosing parameters, the value of each input image  $f$  is scaled inside  $[0, 1]$  by dividing by 255.

### 5.1 Data and shape driven properties of the proposed model

We first consider a synthetic image inside which there are a few different geometric shapes as shown in Fig. 2. In the first row inside this image, geometric shapes include: a diamond, a staircase-like shape, and a square, two of which are convex shaped while the middle one is not convex. These three regions are chosen to have the same area purposely.

First, from the plots, one can see that as the model parameter  $b$  for curvature increases while fixing all the other parameters, objects of smaller sizes are omitted in the final segmentation contour. This suggests a *data driven property* of the proposed model, that is,

the model ignores objects of relatively small sizes while only keeping bigger ones. This feature is also possessed by other models for image denoising, such as the *TV-L1* image denoising model [9] and the mean-curvature denoising model [43]. In fact, it is very natural for the proposed model to assume this feature, since as discussed in Theorem 1, each closed segmentation contour  $\Gamma$  leads to at least  $2\pi$  for the term  $\int_{\Gamma} |\kappa| ds$ , and if the region enclosed by  $\Gamma$  is relatively small, it will be omitted if this causes a change of the fitting terms that is less than  $2\pi$ .

Secondly, from these plots, especially the middle in the second row, one observes that the staircase-like shape is taken out from the final segmentation contour. Notice that the three shapes in the first row of the image share the same area, that is, the fitting terms of the model will treat them equally. Therefore, it is the  $L^1$ -Euler’s elastica term that separates the two convex regions from the non-convex one. By the same argument as above, whether a region is kept in the final segmentation depends on the competition between the fitting terms and the  $L^1$ -Euler’s elastica term. This phenomenon also suggests a new *shape driven property* of the model, that is, it prefers objects of convex shapes. To the best of our knowledge, no discussion about this property has been given for other segmentation models, such as the Mumford-Shah model [30], the Caselles-Kimmel-Sapiro model [7] or the Chan-Vese model [8].

From the depicted results, one may also note that as the parameter  $b$  increases, the contour on the “pac-man” (the right on the second row) becomes more and more convex, since this reduces the  $L^1$ -Euler’s elastica energy. This observation again illustrates the effect of the curvature term in the model.

## 5.2 Experiments for real images and the effect of parameters

In Fig. 3, we apply the model to a real image with a boy wearing a big black hat. This hat is partially occluded by the boy’s face and neck. Image intensity based models, such as the Chan-Vese model, generally find the visible part of the hat with a big notch near the boy’s neck. However, when the proposed segmentation model is applied to this image and with suitable model and Lagrange parameters, this indentation on the segmentation contour can be replaced by an almost straight segment in order to form a convex contour as shown in Fig. 3.

In Fig. 4, we consider another real image inside which there is a mushroom. The mushroom consists of two parts – a stem and a cap, which form a nonconvex shape. By applying the proposed model with a relatively large curvature parameter  $b$ , one notices that only the larger cap part is included in the final segmentation, which cannot be achieved by the conventional intensity based segmentation models.

The above two experiments on real images demonstrate that the proposed segmentation model promotes convex segmentation contours once the curvature parameter  $b$  is chosen large. Moreover, to get those convex contours, the model may automatically restore missing or occluded parts or subtract nonconvex parts.

To see how the curvature parameter  $b$  affects the result of the segmentation, in Fig. 5 and Fig. 6, we present the results for the same two experiments with different values of  $b$  while keeping all the other parameters. From these results, one sees that once  $b$  is small, just as expected, the performance of the new model degenerates to that of the original Chan-Vese’s model. These results also demonstrate that the curvature is captured accurately with the proposed algorithm.

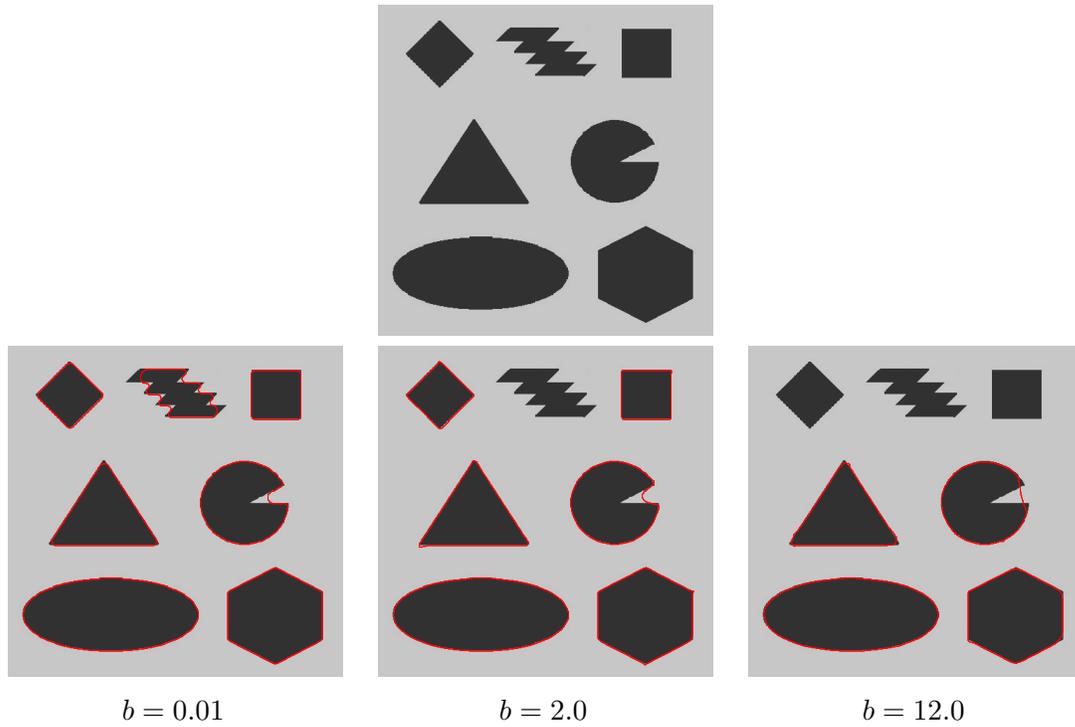


Figure 2: The original image (the first row) and the segmentation results (the second row) with different curvature parameter  $b$ . Two features can be observed from these results: 1) data driven property: as the parameter  $b$  increases, objects of relatively small size will be omitted in the final segmentation; 2) shape driven property: with the same parameter  $b$ , among those objects with equal areas, the one without a convex shape is taken out from the final segmentation (see the staircase-like shape). In this experiment, the other parameters are:  $a = 10^{-4}$ ,  $r_1 = 50$ ,  $r_2 = 20$ ,  $r_3 = 5$ .



Figure 3: The original image and the segmentation result. The result demonstrates that the invisible part of the hat can be restored by the proposed segmentation model with a relatively large curvature parameter  $b$ . In this experiment, the parameters are  $a = 10^{-3}$ ,  $b = 80$ ,  $r_1 = 60$ ,  $r_2 = 40$ ,  $r_3 = 10$ .

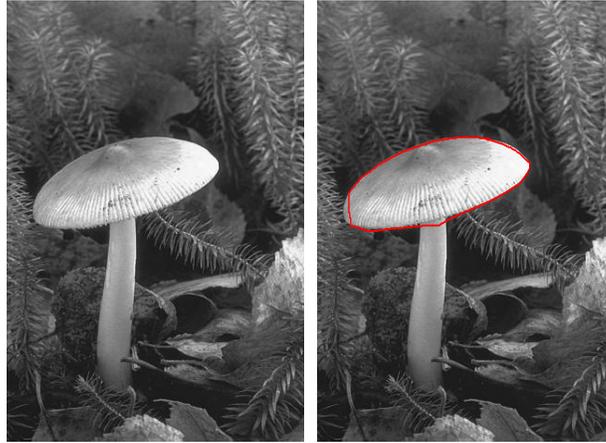


Figure 4: The original image and the segmentation result. The result shows that only the “cap” part of the mushroom is captured while the “stem” is omitted in order to form a convex segmentation contour by the proposed segmentation model with a relatively large curvature parameter  $b$ . In this experiment, the parameters are  $a = 10^{-4}, b = 18, r_1 = 20, r_2 = 10, r_3 = 5$ .



Figure 5: The effect of the curvature parameter  $b$  on the final segmentation results. One may observe that the smaller the parameter  $b$  is chosen, the more salient the indentation of the hat would be. In these two experiments, all the other parameters are the same as those in Fig. 3.

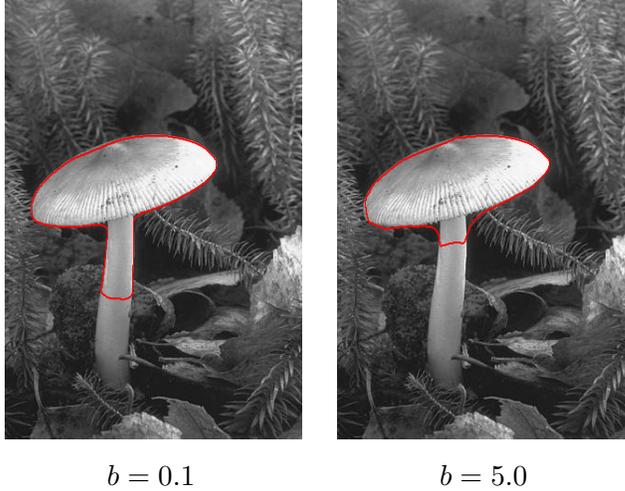


Figure 6: The effect of the curvature parameter  $b$  on the final segmentation results. One can see that once the parameter  $b$  is chosen smaller, larger parts of the stem will be allowed in the final segmentation. In these two experiments, all the other parameters are the same as those in Fig. 4.

To check whether the iteration used for the proposed ALM converges to a saddle point of the augmented Lagrangian functional (Eq. 9), we present the plots of the relative residuals (Eq. 43), relative errors of the Lagrange multipliers (Eq. 44), and relative error of the iterative  $u^k$  (Eq. 45) in Fig. 7. All these plots indicate that a saddle point of the functional and thus a minimizer of the proposed model (Eq. 5) is approached.

### 5.3 Reinitialization of the level set function $\phi$

In the proposed model (Eq. 5), a level set function  $\phi$  is used for locating the segmentation contours. Differently from the alternative model (Eq. 6) with  $u \in [0, 1]$ , no constraint on the value of  $\phi$  is imposed. Hence, when a stationary state of the segmentation contour is attained, that is, the region  $\{\phi > 0\}$  stops evolving, the function  $\phi$  may still keep changing, even though these updates become smaller and smaller. This is because of the choice of the smooth regularized version of the Heaviside function (Eq. 9). In fact, in the region  $\{\phi < 0\}$ , when  $\phi \rightarrow -\infty$ , the term  $\frac{1}{2} + \frac{1}{\pi} \arctan(\frac{\phi}{\epsilon}) \rightarrow 0$ , which decreases the first part of the fitting term. Similarly, in the region  $\{\phi > 0\}$ , the value of  $\phi$  will keep increasing during the iterative process. This phenomenon also explains the slow change of the plots of the relative error of the iterative  $u^k$  in Fig. 7, even though the set  $\{\phi > 0\}$  is invariant just after several hundred of iterations.

As discussed previously, the reason of sticking to a level set function, instead of a binary function, in the model (Eq. 5) is to more accurately capture the mean curvature of segmentation contours. Indeed, since the level set function  $\phi$  presents a few layers of its value around its zero level curve, the curvature of segmentation contour can be represented accurately. Therefore, the above issue that  $\phi$  keeps changing won't affect the performance of the model as long as this changing is not abrupt.

However, if the curvature parameter  $b$  has to be chosen very large for some applications, the evolution of  $\phi$  might become more intractable. The well-known procedure called reini-

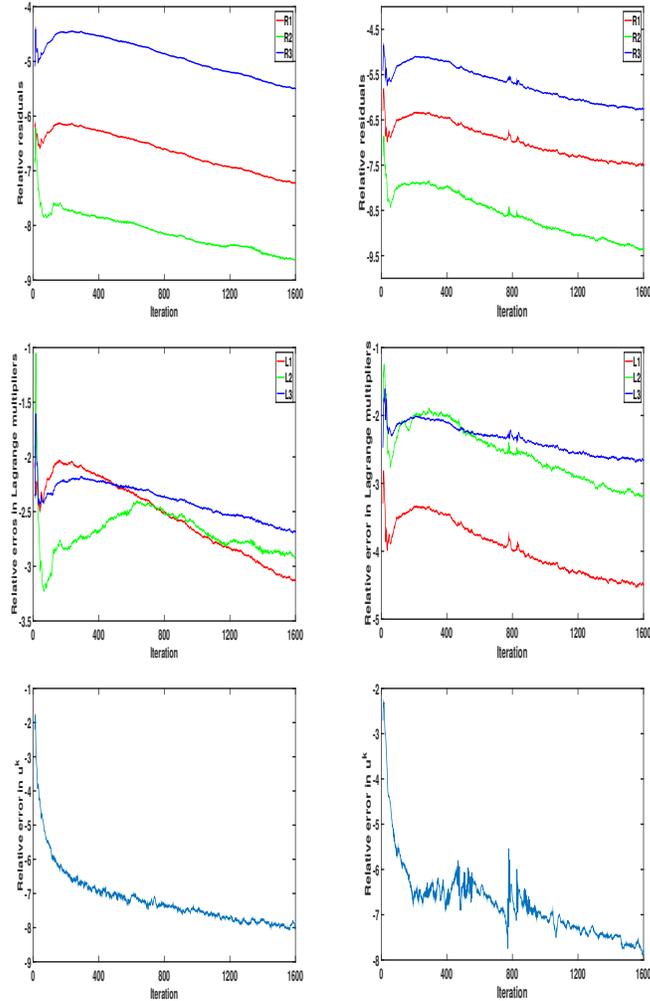


Figure 7: The plots of relative residuals (Eq. 43), relative errors in Lagrange multipliers (Eq. 44), and relative error in  $u^k$  (Eq. 45) for the two examples "hat" (Left column) and "mushroom" (Right column). These plots demonstrate the convergence of the iterative process.

tialization [34] for  $\phi$  is recommended during the iterative process. This will help rationalize the values of  $\phi$  around its zero level set so that the mean curvature can be effectively captured in order to realize the essence of the proposed model.

## 6 Conclusion

This work described a variational image segmentation model, where the  $L^1$ -Euler's elastica energy was used as a regularizer on the segmentation contour. It was shown that an interesting feature of this new model is the ability to promote segmented regions with convex shapes. We also developed a novel augmented Lagrangian method for minimizing the associated functional. Compared with previous work on the  $L^2$ -Euler's elastica energy [39, 44], the proposed algorithm employed fewer Lagrange multipliers, which led to more direct connections among the variables that described curvature and thus helped to express the curvature effectively and accurately. Experiments demonstrated that the new segmentation model could be useful for segmenting objects that were known to have a convex shape. In case occlusions prevented one from observing the entire object, or the transition with the background was not clear by the contrast, meaningful results were still obtained due to the additional convexity information. The experiments also indicated data driven and shape driven properties of the segmentation model, the latter of which has never been discussed in the literature for conventional variational segmentation models to our knowledge.

## References

- [1] L. Ambrosio and S. Masnou, *A direct variational approach to a problem arising in image reconstruction*, Interfaces Free Bound., 5, pp. 63-81, 2003.
- [2] L. Ambrosio and S. Masnou, *On a variational problem arising in image reconstruction*, Vol. 147, Birkhauser, Basel, 2004.
- [3] L. Ambrosio, and V.M. Tortorelli, *Approximation of functionals depending on jumps by elliptic functionals via Gamma convergence*. Comm. Pure Appl. Math. 43, pp. 999-1036, 1990.
- [4] E. Bae and J. Shi and X.-C. Tai, *Graph Cuts for Curvature Based Image Denoising*, IEEE Transactions on Image Processing, 20(5), pp. 1199-1210, 2011.
- [5] C. Brito-Loeza and K. Chen, *Multigrid algorithm for high order denoising*, SIAM J. Imaging Sciences, 3(3), pp. 363-389, 2010.
- [6] K. Bredies, T. Pock, and B. Wirth, *A convex, lower semicontinuous approximation of Euler's elastica energy*, SIAM J. Math. Anal., 7(4), pp. 566-613, 2013.
- [7] V. Caselles, R. Kimmel, and G. Sapiro, *Geodesic active contours*, Int. J. Comput. Vision, 22(1), pp. 61-79, 1997.
- [8] T. Chan and L.A. Vese, *Active contours without edges*, IEEE Trans. Image Process. 10, pp. 266-277, 2001.
- [9] T. Chan and S. Esedoglu, *Aspects of total variation regularized  $L^1$  function approximation*, SIAM J. Appl. Math. 65(5), pp. 1817-1837, 2005.
- [10] T. Chan, S. Esedoglu, and M. Nikolova, *Algorithms for finding global minimizers of denoising and segmentation models*, SIAM J. Appl. Math. 66, pp. 1632-1648, 2006.

- [11] T. Chan, S.H. Kang, and J.H. Shen, *Euler's elastica and curvature based inpaintings*, SIAM J. Appl. Math, Vol. 63(2), pp. 564-592, 2002.
- [12] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser. *Using prior shapes in geometric active contours in a variational framework*. Int. J. Comput. Vision, 50(3), pp: 315-328, 2002.
- [13] D. Cremers, F. Tischhauser, J. Weickert, and C. Schnorr, *Diffusion Snakes: Introducing statistical shape knowledge into the Mumford–Shah functional*, Int. J. Comput. Vision, 50(3), pp. 295-313, 2002.
- [14] T. Chan and W. Zhu, *Level set based shape prior segmentation*, Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1164-1170, 2005.
- [15] E. De Giorgi, *Some remarks on  $\Gamma$ -convergence and least squares methods*, in Composite Media and Homogenization Theory, G. Dal Maso and G.F. Dell' Antonio (Eds.) Birkhauser, Boston, pp. 135-142, 1991.
- [16] M.P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Inc., 1976.
- [17] Y. Duan, Y. Wang, X.-C. Tai, and J. Hahn *A fast augmented Lagrangian method for Euler's elastica model*, SSVM 2011, LSCS 6667, pp.144-156, 2012.
- [18] N. El-Zehiry and L. Grady, *Fast global optimization of curvature*, Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 3257-3264, 2010.
- [19] M. Elsey and S. Esedoglu, *Analogue of the total variation denoising model in the context of geometry processing*, SIAM J. Multiscale Modeling and Simulation Vol. 7(4), pp. 1549-1573, 2009.
- [20] S. Esedoglu and R. March, *Segmentation with depth but without detecting junctions*, J. Math. Imaging and Vision. 18, pp. 7-15, 2003.
- [21] M. Hintermüller, C. N. Rautenberg, and J. Hahn, *Functional-analytic and numerical issues in splitting methods for total variation-based image reconstruction*, Inverse Problems, 30(5), 055014, 2014
- [22] M. Kass, A. Witkin, and D. Terzopoulos, *Snakes: Active contour models*, Int. J. Comput. Vision, 1, pp. 321-331, 1988.
- [23] M. Leventon, W. Grimson, and O. Faugeras, *Statistical shape influence in geodesic active contours*, Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 316-323, 2000.
- [24] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 16, pp. 964-979, 1979.
- [25] R. March and M. Dozio, *A variational method for the recovery of smooth boundaries*, Image and Vision Computing, Vol. 15, pp. 705-712, 1997.
- [26] S. Masnou, *Disocclusion: A variational approach using level lines*, IEEE Trans. Image Process., 11, pp. 68-76, 2002.
- [27] S. Masnou and J.M. Morel, *Level lines based disocclusion*, Proc. IEEE Int. Conf. on Image Processing, Chicago, IL, pp. 259-263, 1998.
- [28] M. Myllykoski, R. Glowinski, T. Kárkkäinen, and T. Rossi, *A new augmented Lagrangian approach for  $L1$ -mean curvature image denoising*, SIAM J. Imaging Sci., 8(1):95-125, 2015.

- [29] E. Strekalovskiy and D. Cremers, *Generalized ordering constraints for multilabel optimization*, IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, 2619–2626, 2011.
- [30] D. Mumford and J. Shah, *Optimal approximation by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., Vol. 42, pp. 577-685, 1989.
- [31] M. Nitzberg, D. Mumford, and T. Shiota, *Filering, segmentation, and depth*, Lecture Notes in Computer Science, Vol. 662, Springer Verlag, Berlin, 1993.
- [32] S. Osher and J. A. Sethian, *Fronts propagating with curvature-dependent speed-algorithm based on Hamilton-Jacobi formulations*, J. Comput. Phys., 79, pp. 12-49, 1988.
- [33] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag, New York, 2002.
- [34] D. Peng, B. Merriman, S. Osher, H.K. Zhao, and M. Kang, *A PDE-based fast local level set method*, J. Comput. Phys, 155, pp. 410-438, 1999.
- [35] R.T. Rockafellar, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Mathematics of Operations Research, Vol 1(2), pp. 97-116, 1976.
- [36] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithm*, Physica D, Vol. 60, pp. 259-268, 1992.
- [37] T. Schoenemann, F. Kahl, and D. Cremers, *Curvature regularity for region-based image segmentation and inpainting: A Linear Programming Relaxation*, IEEE International Conference on Computer Vision (ICCV), 2009.
- [38] T. Schoenemann, F. Kahl, S. Masnou, and D. Cremers, *A linear framework for region-based image segmentation and inpainting involving curvature penalization*, Int. J. Comput. Vision, Volume 99 (1), pp. 53-68, 2012.
- [39] X.C. Tai, J. Hahn, and G.J. Chung, *A fast algorithm for Euler’s Elastica model using augmented Lagrangian method*, SIAM J. Imaging Sciences, Vol. 4(1), pp. 313-344, 2011.
- [40] C. Wu and X.C. Tai, *Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, Vectorial TV, and high order models*, SIAM J. Imaging Sciences, Vol. 3(3), pp. 300-339, 2010.
- [41] F. Yang, K. Chen, and B. Yu, *Homotopy method for a mean curvature-based denoising model*, Appl. Numer. Math., 62(3), pp.185-200, 2012.
- [42] W. Zhu and T. Chan, *A variational model for capturing illusory contours using curvature*, J. Math. Imaging Vision, 27, pp. 29-40, 2007.
- [43] W. Zhu and T. Chan, *Image denoising using mean curvature of image surface*, SIAM J. Imaging Sciences, 5(1), pp. 1-32, 2012.
- [44] W. Zhu, X.C. Tai, and T. Chan, *Image segmentation using Euler’s elastica as the regularization*, J. Sci. Comput., 57(2), pp. 414-438, 2013.
- [45] W. Zhu, X.C. Tai, and T. Chan, *Augmented Lagrangian method for a mean curvature based image denoising model*, Inverse Probl. Imag., 7(4), pp. 1409-1432, 2013.