# *Courbure discrète : théorie et applications*

Rencontre organisée par :
Laurent Najman and Pascal Romon

18-22 novembre 2013

Xue-Cheng Tai

**Fast numerical schemes related to curvature minimization: a brief and elementary review**

Vol. 3, n° 1 (2013), p. 17-30.

<http://acirm.cedram.org/item?id=ACIRM_2013__3_1_17_0>

Centre international de rencontres mathématiques
U.M.S. 822 C.N.R.S./S.M.F.
Luminy (Marseille) France

Actes des rencontres du C.I.R.M.
Vol. **3** n° 1 (2013) 17-30

# Fast numerical schemes related to curvature minimization: a brief and elementary review

Xue-Cheng Tai

**Abstract**

We will treat variational models that use Euler's elastica and related higher order derivatives as regularizers. These models normally lead to higher order partial differential equations with complicated nonlinearities. It is difficult to solve these equations numerically. Recently, some fast numerical techniques have been proposed that can solve these equations with very good numerical speed. We will try to explain the essential ideas of these numerical techniques and point to some central implementation details for these algorithms.

## 1. Introduction

Variational models are becoming essential for image processing and computer vision. A variational model normally needs to minimize an energy functional. This energy functional usually has a "fitting" part and also a "regularizer" part. In this work, we will specially be interested in regularizers involving higher order derivatives.

The goal of image denoising is to remove noise while keeping meaningful vision information such as object edges and boundaries. It is a crucial step in image processing with a wide range of applications in medical image analysis, video monitoring, and others. One of the most popular variational models was proposed by Rudin, Osher, and Fatemi in their seminal work (ROF model) [34]. In [34], a cleaned image is obtained by minimizing the following energy functional

$$E(u) = \int_\Omega |\nabla u| + \frac{\eta}{2}\int_\Omega (f-u)^2, \tag{1.1}$$

where $f : \Omega \to \mathbb{R}$ is a given noisy image defined on $\Omega$, $\int_\Omega |\nabla u|$ stands for the total variation of a function $u$ (see [40] for a definition), and $\eta > 0$ is a positive tuning parameter controlling how much noise will be removed. The remarkable feature of the ROF model lies in its effectiveness in preserving object edges while removing noise. In fact, the total variation regularizer has been widely employed in accomplishing other image processing tasks such as deblurring, segmentation, and registration.

In order to incorporate more geometrical information into the regularizer, a number of higher order regularization models have been proposed and used for image processing and computer vision problems. In this work, we will mainly consider three higher models and outline some fast algorithms to solve them, c.f. Section 2. To make the presentation clearer, we will only state these models for simple image restoration problems with a given noisy image $f$. There exist extensions of these models for more general applications related to image processing and computer vision including deblurring, inpainting, zooming and geometry minimization.

---

Text presented during the meeting "Discrete curvature: Theory and applications" organized by Laurent Najman and Pascal Romon. 18-22 novembre 2013, C.I.R.M. (Luminy).
2000 *Mathematics Subject Classification.* 00X99.
*Key words.* variaitonal models, curvature minimization, Augmented Lagrangian methods.
Supported by the Christian Michelsen Research (CMR), Bergen.

## 2. Higher order regularizations

The ROF model has several unfavorable features. The main caveat is the stair-case effect, that is, the resulting clean image would present blocks even though the desired image could be smooth. Other undesirable properties include corner smearing and loss of image contrast. To remedy these drawbacks, a very rich list of results exists in the literature, see [24, 1, 11, 49, 52]. Despite the effectiveness of these models in removing the staircase effect, it is often a challenging issue to minimize the corresponding functionals. Note that the models contain second order derivatives, the related Euler-Lagrange equations are fourth-order. It is a nontrivial task to develop effective and efficient algorithms to solve these higher order nonlinear equations.

In this section, we will first give an informal introduction to three higher order models. In the subsequent sections, we will introduce fast algorithms to solve them.

**Remark 1.** We remark on a few important issues for the models and algorithms:

(1) The introduction about the higher order mathematical models is very informal in a mathematical sense. To define and analyze these models for proper function spaces is far more difficult and is beyond the content of this note. Analysis of these models in the continuous setting is still incomplete in the literature.

(2) For image processing problems, the computation domain is always a rectangle. The pixels of an image give a ready mesh for the discretization. This leads to easy discretization with finite difference approximations for our models and algorithms. However, this is not a restriction for the models. The models and algorithms are valid for general domains as well. For general domains with curved boundaries, the discrete approximation of the functions (including primal and auxiliary functions) could be complicated near the boundaries.

(3) In the literature, discrete curvature is often used, see other publications in this proceeding. Instead, we have chosen to present these models in a continuous setting. The algorithms are also presented in the continuous setting. For numerical implementation, we must discretize the continuous functions.

Both approaches ("discrete curvature" or "continuous curvature") have advantages and disadvantages. For many discrete curvature models, it seems that the "discrete curvature" is not converging to the "continuous curvature" when the mesh size goes to zero. For the continuous models, it is easy to see that we are using the curvature. Standard finite difference or finite element approximations would lead to natural approximations for the curvature terms. However, a rigorous proof for this is still missing due to the complexity of the models.

### 2.1. Regularization using TV2.

In [24], Lysaker et al. directly incorporated second order derivative information into the image denoising process. They proposed to minimize the following energy functional to denoise an image:

$$E(u) = \int_\Omega \sqrt{u_{xx}^2 + u_{xy}^2 + u_{yx}^2 + u_{yy}^2} + \frac{\eta}{2}\int_\Omega (f-u)^2. \tag{2.1}$$

This higher order energy functional is much simpler than the elastic regularizer that we shall introduce later. Numerically, this regularizer shows rather good performance with noise suppression and edge preservation. In the literature, there exists quite a number of related models, see [13, 20, 42, 3, 12, 6, 18, 36, 31, 5, 14, 8, 19, 48, 30, 7]. The well-posedness for this energy functional and its gradient flow equation have been studied in [27, 26, 17].

### 2.2. Regularization using the Euler's Elastica energy.

Given a function $f : \Omega \mapsto R$, the Euler's elastica model needs to find a function $u$ to minimize the following energy functional:

$$E(u) = \int_\Omega \left[a + b\left(\nabla \cdot \frac{\nabla u}{|\nabla u|}\right)^2\right] |\nabla u| + \frac{\eta}{2}\int_\Omega (f-u)^2. \tag{2.2}$$

The nonnegative constants $a, b$ and $\eta$ need to be chosen properly for different purposes of applications. This model comes from the Euler's Elastca energy for curves, see [11] for some more explanations for this energy. For a given curve $\Gamma \subset R^2$ with curvature $\kappa$, the Euler's elastica

energy is defined as:

$$\int_\Gamma (a + b\kappa^2) ds.$$

For a function $u$, the curvature for the level curve $\Gamma_c : u(x) = c$ is:

$$\kappa = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right).$$

Thus, the Euler's elastica energy for the level curve: $u(x) = c$ is:

$$\ell(c) = \int_{\Gamma_c} \left( a + b \left[ \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \right]^2 \right) ds.$$

Summing up (integrating) the Euler's elastica energy for all the level curves $\Gamma_c : c \in (-\infty, \infty)$, we get from the co-area formula [39] that the total Euler's elastica energy for all the level curves is:

$$\int_{-\infty}^{\infty} \ell(c) dc = \int_{-\infty}^{\infty} \int_{\Gamma_c} \left( a + b \left[ \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \right]^2 \right) ds dc = \int_\Omega \left( a + b \left[ \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \right]^2 \right) |\nabla u| dx.$$

Minimization problem (2.2) is trying to use the total Euler's Elastica energy as a regularizer to remove noise from the image $f$.

2.3. **Regularization using the image surface mean curvature.** In [49], the authors proposed a variational model that uses the mean curvature of the induced image surface $(x, y, f(x, y))$ to remove noise. Specifically, the model employs the $L^1$ norm of the mean curvature of the image surface as the regularizer and minimizes the following functional to get a clean image:

$$E(u) = \int_\Omega \left| \nabla \cdot \left( \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) \right| + \frac{\eta}{2} \int_\Omega (f - u)^2. \tag{2.3}$$

Above, $\eta$ is a tuning parameter. The term $\nabla \cdot (\frac{\nabla u}{\sqrt{1+|\nabla u|^2}})$ is the mean curvature of the surface $\phi(x, y, z) = u(x, y) - z = 0$. The model tries to fit the given noisy image surface $(x, y, f(x, y))$ with a surface $(x, y, u(x, y))$ that is regularized by the mean curvature. This idea can be traced to much earlier papers, see [22]. The model can sweep noise while keeping object edges, and it also avoids the staircase effect. More importantly, as discussed in [49, 53], the model is also capable of preserving image contrasts as well as object corners.

## 3. Fast numerical algorithms based on augmented Lagrangian method (ALM)

In this section, we first show the split-Bregman altorithm of Goldstein-Osher [16] for the ROF model [34]. We then extend this idea for the three higher order models we have introduced in Section 2 to get fast algorithms for them.

3.1. **Split-Bregman for ROF.** In work by Goldstein-Osher [16], fast iterative schemes were proposed and tested for the ROF model. It is one of the most efficient numerical schemes for solving the ROF model. Later, it was observed that the split-Bregman algorithm of Golstein-Osher [16] is equivelent to the Augmented Lagrangian method [38, 41]. Here, we explain the ideas in an "informal" fashion. We will present the ideas in a continuous setting. As stated in Remark 1, to make our statements precise, more precise definitions of the function spaces and the norms need to be given. That is one of the reasons that discrete models have been used to explain these algorithms as in [41]. We will not get into the details related to this kind of technicalities.

Let $p = \nabla u$, then it is easy to see that the ROF model is equivalent to the following constrained minmization problem:

$$\min_{\substack{u,p \\ p=\nabla u}} \int_\Omega |p| + \frac{\eta}{2} |u - f|^2 dx. \tag{3.1}$$

Let us use the Augmented Lagrangian method (ALM) [15] to deal with the constraint $p = \nabla u$. Then we need to find a saddle point for the following Lagrangian functional:

$$L_{rof}(u, p, \lambda) = \int_\Omega |p| + \frac{\eta}{2} |u - f|^2 + \lambda \cdot (p - \nabla u) + \frac{r}{2} |p - \nabla u|^2 dx. \tag{3.2}$$

Above: $u : \Omega \mapsto R$ denotes the image we need to find, $p : \Omega \mapsto R^2$ is a vector valued function related to the gradient of the function $u$, $\lambda : \Omega \mapsto R^2$ denotes the Lagrangian multiplier. Due to convexity, problem (3.1) has a unique solution (in the discrete setting as well as in a proper space in the continuous setting). If $(u^*, p^*)$ is a global minimizer for (3.1), then there exists a $\lambda^*$ such that $(u^*, p^*, \lambda^*)$ is a saddle point for (3.2). We propose to use Algorithm 1 to search for a saddle point for (3.2).

**Algorithm 1** Augmented Lagrangian method for the ROF model

Initialization: $\lambda^0 = 0, u^0 = f$; For k=0,1,2,...:

(1) Compute $p^{k+1}$ from :

$$p^{k+1} = \arg\min_q L_{rof}(u^k, q; \lambda^k), \tag{3.3}$$

(2) Compute $u^{k+1}$ from:

$$u^{k+1} = \arg\min_v L_{rof}(v, p^{k+1}; \lambda^k), \tag{3.4}$$

(3) Update

$$\lambda^{k+1} = \lambda^k + r(p^{k+1} - \nabla u^{k+1}). \tag{3.5}$$

(4) Go to the next iteration if not converged.

Minimization subproblem (3.3) has closed-form solutions and thus can be easily computed. Minimization subproblem (3.4) can be solved by FFT (Fast Fourier Transform) or simple Gauss-Seidel iterative solvers to get an approximate solution. See [43, 41] for more details. Theoretically, it is necessary to have sufficiently many iterations between subproblems (3.3) and (3.4). In practice, the above algorithm works well for most of the cases for the ROF model.

It is also easy to extend the above model for vector-valued functions and vector-TV regularization, see [41, p.320] and [33, 32].

3.2. **Split-Bregman for second order Total variation.** Here, we explain how to use the fast algorithm explained in the last section for the regularization model (2.1) related to second order derivatives. The idea follows the work [41].

The essential idea for the fast schemes is to introduce some auxiliary variables and consider the complicated minimization problem as a constrained minimization. We then use splitting ideas to decompose the complicated minimization problem into some simpler minimization problems. Let

$$p = D^2 u = \begin{pmatrix} u_{xx} & u_{xy} \\ u_{yx} & u_{yy} \end{pmatrix}.$$

Thus $p$ is a matrix function defined on $\Omega$, i.e. $p$ is equal to the Hessian matrix of $u$ over $\Omega$. The minimization of the energy functional given in (2.1) is equivalent to:

$$\min_{\substack{u,p \\ p=D^2u}} \int_\Omega |p| + \frac{\eta}{2}|u - f|^2 dx. \tag{3.6}$$

Above, $|p| = \sqrt{\sum_{i,j} p_{ij}^2}$ stands for the Frobenius norm of the matrix $p$.

Again, we use the Augmented Lagrangian method (ALM) [15, 41] to deal with the constraint $p = D^2 u$. Then we need to find a saddle point for the following Lagrangian functional:

$$L_{llt}(u, p, \lambda) = \int_\Omega |p| + \frac{\eta}{2}|u - f|^2 + \lambda : (p - D^2 u) + \frac{r}{2}|p - D^2 u|^2 dx. \tag{3.7}$$

Here $u : \Omega \mapsto R$ denotes the image we need to find, $p : \Omega \mapsto R^4$ is a matrix valued function related to the Hessian of the function $u$, $\lambda : \Omega \mapsto R^4$ denotes the Lagrangian multiplier. The notation $A : B$ denotes the elementwise inner product of two matrices $A$ and $B$. We use Algorithm 2 to search for a saddle point for (3.7).

Similar to Algorithm 1, the minimization subproblem (3.8) needs to compute a matrix-valued function and it has closed-form solutions and thus can be easily computed. Minimization subproblem (3.9) gives raise to a linear 4th order equation on a regular mesh. It can be solved by FFT

**Algorithm 2** Augmented Lagrangian method for the TV2 model

*Initialization*: $\lambda^0 = 0, u^0 = f$; *For k=0,1,2,...*:

(1) Compute $p^{k+1}$ from

$$p^{k+1} = \arg\min_q L_{llt}(u^k, q; \lambda^k), \tag{3.8}$$

(2) Compute $u^{k+1}$ from:

$$u^{k+1} = \arg\min_v L_{llt}(v, p^{k+1}; \lambda^k), \tag{3.9}$$

(3) Update

$$\lambda^{k+1} = \lambda^k + r(p^{k+1} - D^2 u^{k+1}). \tag{3.10}$$

(4) Go to the next iteration if not converged.

(Fast Fourier Transform) or simple Gauss-Seidel iterative solvers to get an approximate solution. See [41, p.324] for more details.

3.3. **Augmented Lagrangian method for Euler's elastica model.** In order to use fast algorithms related to ALM for the minimization of the Euler's elastica given in (2.2), it is necessary to introduce a few more auxiliary functions. The ideas presented here follow the work [37]. The following lemma is easy to prove using Hölder's inequality:

**Lemma 2.** *Let $n \neq 0$ and $p \neq 0$ be two given vectors. They satisfy*

$$|n| \leq 1, |p| = n \cdot p,$$

*if and only if*

$$n = \frac{p}{|p|}.$$

Let us define

$$p = \nabla u, \quad n = \frac{p}{|p|}.$$

It is easy to see that the minimization of the Euler's elastica energy (2.2) is equivalent to the following constrained minimization:

$$\begin{aligned} &\min_{u,p,n} \int_\Omega \left(a + b(\nabla \cdot n)^2\right)|p| + \frac{\eta}{2}\int_\Omega |u - f|^2 \\ &\text{with} \quad p = \nabla u, \quad |p| = n \cdot p, \quad |n| \leq 1. \end{aligned} \tag{3.11}$$

The use of $n$ with $|n| \leq 1$ can be viewed as a relaxation. Moreover, the constraint $|n| \leq 1$ is crucial to prevent the unboundedness of $n$ when $p = \mathbf{0}$. Define the characteristic function $\delta_{\mathcal{R}}(\cdot)$ on $\mathcal{R}$ as

$$\delta_{\mathcal{R}}(m) = \begin{cases} 0 & m \in \mathcal{R} \equiv \{m \in \mathbf{L}^2(\Omega) \mid |m| \leq 1 \text{ a.e. in } \Omega\}, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.12}$$

then, the constrained minimization problem (3.11) can be rewritten as:

$$\begin{aligned} &\min_{u,p,n} \int_\Omega \left(a + b(\nabla \cdot n)^2\right)|p| + \frac{\eta}{2}\int_\Omega |u - f|^2 + \delta_{\mathcal{R}}(n) \\ &\text{with} \quad p = \nabla u, \quad |p| = n \cdot p, \end{aligned} \tag{3.13}$$

We know that $|n| \leq 1$ in $\Omega$, thus

$$|p| - n \cdot p \geq 0 \quad \forall x \in \Omega.$$

There are two constraints in (3.13). Two different penalizations are used for these two constraints. For constraint $p = \nabla u$, we use the L2-norm for the penalization term; but for $|p| = n \cdot p$, we use

the L1-norm. With this special treatment for the last constraint, the corresponding Lagrangian functional becomes:

$$\begin{aligned} L_{elas}(u,p,n,\lambda_1,\lambda_2) &= \int_\Omega \left(a+b(\nabla\cdot n)^2\right)|p| + \frac{\eta}{2}|u-f|^2 dx + \delta_{\mathcal{R}}(n) \\ &+ \int_\Omega \lambda_1\cdot(p-\nabla u) + \frac{r_1}{2}|p-\nabla u|^2 + \lambda_2(|p|-n\cdot p) + r_2(|p|-n\cdot p)dx. \end{aligned} \tag{3.14}$$

The meanings of the primal and dual variables are listed in the following:

- $u:\Omega\mapsto R$ denotes the image we need to find,
- $p:\Omega\mapsto R^2$ is a vector-valued function related to the gradient of the function $u$,
- $n:\Omega\mapsto R^2$ is a vector-valued function related to the unit vectors of the level curves of $u$,
- $\lambda_1:\Omega\mapsto R^2$ denotes the Lagrangian multiplier for constraint $p=\nabla u$,
- $\lambda_2:\Omega\mapsto R$ denotes the Lagrangian multiplier for constraint $|p|=n\cdot p$.

We shall use Algorithm 3 to search for a saddle point of this Lagrangian functional.

**Algorithm 3** Augmented Lagrangian method for the Euler's elastica model

*Initialization*: $\lambda^0=0, u^0=f, n^0=0$; *For k=0,1,2,...*:

(1) Compute $p^{k+1}$ from

$$p^{k+1} = \arg\min_q L_{elas}(u^k, q, n^k; \lambda^k), \tag{3.15}$$

(2) Compute $n^{k+1}$ from

$$n^{k+1} = \arg\min_m L_{elas}(u^k, p^{k+1}, m; \lambda^k), \tag{3.16}$$

(3) Compute $u^{k+1}$ from:

$$u^{k+1} = \arg\min_v L_{elas}(v, p^{k+1}, n^{k+1}; \lambda^k), \tag{3.17}$$

(4) Update

$$\lambda_1^{k+1} = \lambda_1^k + r_1(p^{k+1}-\nabla u^{k+1}), \tag{3.18}$$

$$\lambda_2^{k+1} = \lambda_2^k + r_2(|p^{k+1}| - n^{k+1}\cdot p^{k+1}). \tag{3.19}$$

(5) Go to the next iteration if not converged.

In the following, we give some remarks on the solutions of the subproblems and some implementation issues for Algorithm 3:

(1) Minimization subproblem (3.15) has closed-form solutions. A simple thresholding is sufficient to get the solution $p^{k+1}$.

(2) Minimization subproblem (3.16) can be approximated by solving a linear partial differential equation first and then projecting the obtained solution onto the convex set $\mathcal{R}$ defined in (3.12). The linear equation is:

$$-b\nabla(|p^{k+1}|\nabla\cdot n^{k+1}) = (\lambda_2+r_2)p^{k+1}. \tag{3.20}$$

This equation has some similarities with the gradient-divergence equations from the Maxwell equation for magnetic simulations or Darcy-Stokes flow, see [25]. It can be approximately solved by a few Gauss-Seidel iterations or a coefficient freezing FFT solver, see [37] for more details. Note that both $p$ and $n$ are vector-valued functions. We need to solve for all the components of the vector functions.

(3) The minimizer $u^{k+1}$ of subproblem (3.17) satisfies the following linear partial differential equation:

$$u^{k+1} - f + \nabla\cdot\lambda_1^k + r_1\nabla\cdot(p^{k+1}-\nabla u^{k+1}) = 0.$$

This equation is the same as for the ROF model, c.f. Algorithm 1. It can be easily solved by FFT or few Gauss-Seidel iterations.

(4) Theoretically, we need sufficiently many iterations between (3.15)-(3.17) to guarantee convergence of the algorithm. In practice, just one iteration, as stated in Algorithm 3, is enough to have convergence of the iterative solutions.

(5) As the energy functional is not convex, we need to choose some of the penalization parameters $r_i$ sufficiently big to obtain convergence of this algorithm. Tuning the parameters $r_i$ is a delicate issue. Fortunately, there exists an easy way to get the correct values for these penalization parameters $r_i$. We outline the details in section 5.
(6) Another relaxation method for the Euler's Elastica model was proposed in [4]. It solves the problem by means of tractable convex relaxation in higher dimensions.

The algorithm presented in [37] for the minimization of the Euler's elastica energy has one more auxiliary function variable. There, the following constrained minimization was considered:

$$\begin{aligned} &\min_{u,p,n} \int_\Omega \left(a + b(\nabla \cdot n)^2\right)|p| + \frac{\eta}{s}\int_\Omega |u-f|^s + \delta_{\mathcal{R}}(m) \\ &\text{with} \quad p = \nabla u, \quad |p| = n \cdot p, \quad n = m, \end{aligned} \tag{3.21}$$

Note that the fidelity term has also been changed and it is powered by $s$. Normally, the value of $s$ can be chosen as $s = 1$ or $s = 2$ depending on the nature of the noise contained in $f$. For salt-and-pepper noise, we prefer to choose $s = 1$. For Gaussian noise, we choose $s = 2$. There could exist cases where we need to choose $s \in [1, \infty]$.

In [37], the following Lagrangian functional is used for the above constrained minimization problem (3.21):

(3.22)

$$\begin{aligned} L_{elas}(u,p,m,n,\lambda_1,\lambda_3,\lambda_3) &= \int_\Omega \left(a + b(\nabla \cdot n)^2\right)|p| + \frac{\eta}{s}|u-f|^s dx + \delta_{\mathcal{R}}(m) \\ &+ \lambda_1 \cdot (p - \nabla u) + \frac{r_1}{2}|p - \nabla u|^2 + \lambda_2(|p| - n \cdot p) + r_2(|p| - n \cdot p) + \lambda_3 \cdot (m-n) + \frac{r_3}{2}|n-m|^2 dx \end{aligned}$$

An algorithm similar to Algorithm 3 can be used to find a saddle point for the above Lagrangian functional. We will not repeat the details. We can see that the minimization subproblem for $n$ does not have the constraint $|n| \leq 1$ and we only need to solve a linear PDE system to get the values of $n^{k+1}$ which can be done by using FFT or a few Gauss-Seidel iterations, c.f. (3.20). The convex constraint is only imposed on $m$ now. The solution of the minimization subproblem for $m$ is in fact just a simple projection to the convex set, see [37] for the details.

It is also possible to use these ideas for a generalized Euler's elastica model with the energy functional modified to be:

$$E(u) \quad = \quad \int_\Omega \left[a + b\left|\left(\nabla \cdot \frac{\nabla u}{|\nabla u|}\right)\right|^{s_1}\right]|\nabla u| + \frac{\eta}{s_2}\int_\Omega |f-u|^{s_2}. \tag{3.23}$$

In case that $s_1 = 2, s_2 = 1$, we could consider the following splitting:

$$\begin{aligned} &\min_{u,p,n} \int_\Omega \left(a + b(\nabla \cdot n)^2\right)|p| + \eta\int_\Omega |v-f| + \delta_{\mathcal{R}}(n) \\ &\text{with} \quad v = u \quad p = \nabla u, \quad |p| = n \cdot p, \end{aligned} \tag{3.24}$$

It is easy to define the corresponding Lagrangian functional and use an alternating minimization scheme to search for its saddle point. The details of the corresponding algorithm will be omitted and all the minimization subproblems can be easily solved or have closed-form solutions, see [37, p.33] for the needed details.

If we consider the case that $s_1 = 1, s_2 = 2$, then it would be better to use the following splitting idea:

$$\begin{aligned} &\min_{u,p,n} \int_\Omega (a + b|q|)|p| + \frac{\eta}{2}\int_\Omega |u-f|^2 + \delta_{\mathcal{R}}(m) \\ &\text{with} \quad q = \nabla \cdot n, \quad p = \nabla u, \quad |p| = n \cdot p, \quad n = m. \end{aligned} \tag{3.25}$$

It is easy to define the corresponding Lagrangian functional and use an alternating minimization scheme to search for its saddle point. The details of the corresponding algorithm will be omitted and all the minimization subproblems can be easily computed or have closed-form solutions, see [37, p.33] for the needed details.

We want to emphasise that for the constraint $|p| - m \cdot p = 0$, we use L1-norm for the penalization and it is true that $|p| - m \cdot p \geq 0$ due to the fact that $|m| \leq 1$.

So far, we have explained the splitting techniques needed for the cases $s_1 = 2, s_2 = 1$ and $s_1 = 1, s_2 = 2$. For the case $s_1 = s_2 = 1$, we need to combine the splitting techniques for both cases. These techniques can be easily extended for more general values of $s_1$ an $s_2$. In fact, the algorithm given in [37] can deal with general $Lp$ fidelity terms.

3.4. **Augmented Lagrangian method for the mean curvature (MC) model.** The ideas presented in this section follow the work [53]. Let us recall the idea of introducing the mean curvature denoising model. In this model, a $2D$ image $f(x, y)$ is regarded as a surface $(x, y, f(x, y))$ in $\mathbb{R}^3$, c.f. [22, 49]. One thus considers the surface $\phi(x, y, z) = u(x, y) - z = 0$ and the mean curvature $\kappa = \nabla \cdot (\nabla\phi/|\nabla\phi|) = \nabla \cdot (\langle \nabla u, -1\rangle/|\langle \nabla u, -1\rangle|)$. Here and later, $\langle \cdot, \cdot \rangle$ is used to denote the concatenation of vectors. Note that one introduces two variables $\mathbf{p} = \nabla u$ and $\mathbf{n} = \nabla u/|\nabla u|$ to tackle the Euler's elastica for its curvature term $\kappa = \nabla \cdot (\nabla u/|\nabla u|)$. This gives us a hint on how to treat the curvature term in our case, that is, we may introduce a variable $\mathbf{p} = \langle \nabla u, -1\rangle$ instead of $\mathbf{p} = \nabla u$. Accordinlgy, we will also introduce $\mathbf{n} = \langle \nabla u, -1\rangle/|\langle \nabla u, -1\rangle|$.

With constraints

$$\mathbf{p} = \langle \nabla u, -1\rangle, \mathbf{n} = \langle \nabla u, -1\rangle/|\langle \nabla u, -1\rangle|,$$

the MC model (2.3) is then transformed to the following constrained minimization problem:

$$\begin{aligned} &\min_{u,q,\mathbf{n},\mathbf{p}} \left[\lambda \int_\Omega |q| + \frac{1}{2}\int_\Omega (f-u)^2\right], \\ &\text{with } q = \nabla \cdot \mathbf{n},\ \mathbf{n} = \frac{\mathbf{p}}{|\mathbf{p}|},\ \mathbf{p} = \langle \nabla u, -1\rangle. \end{aligned} \tag{3.26}$$

The associated augmented Lagrangian functional is then:

$$\begin{aligned} \mathcal{L}(u, q, \mathbf{p}, \mathbf{n}, \mathbf{m}; \lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_4) &= \lambda \int_\Omega |q| + \frac{1}{2}\int_\Omega (f-u)^2 \\ &+ r_1 \int_\Omega (|\mathbf{p}| - \mathbf{p}\cdot\mathbf{m}) + \int_\Omega \lambda_1(|\mathbf{p}| - \mathbf{p}\cdot\mathbf{m}) \\ &+ \frac{r_2}{2}\int_\Omega |\mathbf{p} - \langle \nabla u, -1\rangle|^2 + \int_\Omega \boldsymbol{\lambda}_2 \cdot (\mathbf{p} - \langle \nabla u, -1\rangle) \\ &+ \frac{r_3}{2}\int_\Omega (q - \partial_x n_1 - \partial_y n_2)^2 + \int_\Omega \lambda_3 (q - \partial_x n_1 - \partial_y n_2) \\ &+ \frac{r_4}{2}\int_\Omega |\mathbf{n} - \mathbf{m}|^2 + \int_\Omega \boldsymbol{\lambda}_4 \cdot (\mathbf{n} - \mathbf{m}) + \delta_{\mathcal{R}}(\mathbf{m}), \end{aligned} \tag{3.27}$$

where $r_i's$, $i = 1, \cdots, 4$, are the penalization parameters, and $\lambda_1, \lambda_3 : \Omega \mapsto \mathbb{R}$ and $\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_4 : \Omega \mapsto \mathbb{R}^3$ are Lagrange multipliers, and $\mathbf{p}, \mathbf{n}, \mathbf{m} : \Omega \mapsto \mathbb{R}^3$ are vector-valued functions. For the sake of completeness of presentation, we make a few remarks in the following.

Introduction of the variable $\mathbf{m}$ aims to relax variable $\mathbf{n}$ that is defined as $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$. The variable $\mathbf{m}$ is required to lie in the set $\mathcal{R}$ so that the term $|\mathbf{p}| - \mathbf{p}\cdot\mathbf{m}$ is always non-negative. As discussed in [37], the benefit of this non-negativeness is that the $L^2$ penalization is unnecessary. Instead, we use L1-norm for the penalization. As this term is always positive, the penalization term becomes just $|\mathbf{p}| - \mathbf{p}\cdot\mathbf{m}$.

As the saddle points of the augmented Lagrangian functional (3.27) correspond to the minimizers of the constrained minimization problem (3.26), one just needs to find the saddle points of (3.27). Similar to algorithms for the Euler's elastica model, we apply an iterative procedure. Specifically, for each variable in (3.27), we fix all the other variables and seek a critical point of the induced functional to update this variable. Once all the variables are updated, the Lagrangian multipliers will also be updated. Then we repeat the process until the variables converge to a steady state. The algorithm is summarized in Algorithm 4.

The sub-minimization problems (3.28)-(3.32) are very easy to solve. We list their corresponding minimization energy functionals in the following:

$$\varepsilon_1(u) = \frac{1}{2}\int_\Omega (f-u)^2 + \frac{r_2}{2}\int_\Omega |\mathbf{p} - \langle \nabla u, -1\rangle|^2 + \int_\Omega \boldsymbol{\lambda}_2 \cdot (\mathbf{p} - \langle \nabla u, -1\rangle), \tag{3.33}$$

$$\varepsilon_2(q) = \lambda \int_\Omega |q| + \frac{r_3}{2}\int_\Omega (q - \partial_x n_1 - \partial_y n_2)^2 + \int_\Omega \lambda_3 (q - \partial_x n_1 - \partial_y n_2), \tag{3.34}$$

**Algorithm 4** Alternating minimization method for surface mean curvature minimization.

(1) Initialization: $u^0$, $q^0$, $\mathbf{p}^0$, $\mathbf{n}^0$, $\mathbf{m}^0$, and $\lambda_1^0$, $\boldsymbol{\lambda}_2^0$, $\lambda_3^0$, $\boldsymbol{\lambda}_4^0$. For $k \geq 1$, loop over the following two steps:

(2) Compute an approximate minimizer $(u^k, q^k, \mathbf{p}^k, \mathbf{n}^k, \mathbf{m}^k)$ of the augmented Lagrangian functional with the fixed Lagrangian multiplier $\lambda_1^{k-1}$, $\boldsymbol{\lambda}_2^{k-1}$, $\lambda_3^{k-1}$, $\boldsymbol{\lambda}_4^{k-1}$ from the following minimization problems:

$$u^k = \operatorname{argmin} \mathcal{L}(u, q^{k-1}, \mathbf{p}^{k-1}, \mathbf{m}^{k-1}, \mathbf{n}^{k-1}, \lambda_1^{k-1}, \boldsymbol{\lambda}_2^{k-1}, \lambda_3^{k-1}, {\boldsymbol{\lambda}_4}^{k-1}) \tag{3.28}$$

$$q^k = \operatorname{argmin} \mathcal{L}(u^k, q, \mathbf{p}^{k-1}, \mathbf{m}^{k-1}, \mathbf{n}^{k-1}, \lambda_1^{k-1}, \boldsymbol{\lambda}_2^{k-1}, \lambda_3^{k-1}, {\boldsymbol{\lambda}_4}^{k-1}) \tag{3.29}$$

$$p^k = \operatorname{argmin} \mathcal{L}(u^k, q^k, \mathbf{p}, \mathbf{m}^{k-1}, \mathbf{n}^{k-1}, \lambda_1^{k-1}, \boldsymbol{\lambda}_2^{k-1}, \lambda_3^{k-1}, {\boldsymbol{\lambda}_4}^{k-1}) \tag{3.30}$$

$$m^k = \operatorname{argmin} \mathcal{L}(u^k, q^k, \mathbf{p}^k, \mathbf{m}, \mathbf{n}^{k-1}, \lambda_1^{k-1}, \boldsymbol{\lambda}_2^{k-1}, \lambda_3^{k-1}, {\boldsymbol{\lambda}_4}^{k-1}) \tag{3.31}$$

$$n^k = \operatorname{argmin} \mathcal{L}(u^k, q^k, \mathbf{p}^k, \mathbf{m}^k, \mathbf{n}, \lambda_1^{k-1}, \boldsymbol{\lambda}_2^{k-1}, \lambda_3^{k-1}, {\boldsymbol{\lambda}_4}^{k-1}) \tag{3.32}$$

(3) Update the Lagrangian multipliers

$$\begin{aligned} \lambda_1^k &= \lambda_1^{k-1} + r_1(|\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}^k) \\ \boldsymbol{\lambda}_2^k &= \boldsymbol{\lambda}_2^{k-1} + r_2(|\mathbf{p}|^k - \langle \nabla u^k, -1 \rangle) \\ \lambda_3^k &= \lambda_3^{k-1} + r_3(q^k - \partial_x n_1^k - \partial_y n_2^k) \\ \boldsymbol{\lambda}_4^k &= \boldsymbol{\lambda}_4^{k-1} + r_4(\mathbf{n}^k - \mathbf{m}^k), \end{aligned}$$

where $\mathbf{n} = \langle n_1, n_2, n_3 \rangle$.

(4) Stop if the given stopping criteria have been satisfied. Otherwise, go to the next iteration.

$$\begin{aligned} \varepsilon_3(\mathbf{p}) &= r_1 \int_\Omega (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \int_\Omega \lambda_1(|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \frac{r_2}{2} \int_\Omega |\mathbf{p} - \langle \nabla u, -1 \rangle|^2 \\ &\quad + \int_\Omega \boldsymbol{\lambda}_2 \cdot (\mathbf{p} - \langle \nabla u, -1 \rangle), \end{aligned} \tag{3.35}$$

$$\begin{aligned} \varepsilon_4(\mathbf{m}) &= r_1 \int_\Omega (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \int_\Omega \lambda_1(|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \frac{r_4}{2} \int_\Omega |\mathbf{n} - \mathbf{m}|^2 \\ &\quad + \int_\Omega \boldsymbol{\lambda}_4 \cdot (\mathbf{n} - \mathbf{m}) + \delta_{\mathcal{R}}(\mathbf{m}), \end{aligned} \tag{3.36}$$

$$\begin{aligned} \varepsilon_5(\mathbf{n}) &= \frac{r_3}{2} \int_\Omega (q - \partial_x n_1 - \partial_y n_2)^2 + \int_\Omega \lambda_3(q - \partial_x n_1 - \partial_y n_2) + \frac{r_4}{2} \int_\Omega |\mathbf{n} - \mathbf{m}|^2 \\ &\quad + \int_\Omega \boldsymbol{\lambda}_4 \cdot (\mathbf{n} - \mathbf{m}). \end{aligned} \tag{3.37}$$

Fast solvers and closed-form solutions are available for all these subproblems, c.f. [53].

## 4. Euler's elastica regularizer for interface problems

The classical snake and active contour model was given by Kass, Witkin, and Terzopoulos [21] where they proposed minimizing the functional

$$\mathbf{E}(\mathcal{C}) = \alpha \int_0^1 |\mathcal{C}'(s)|^2 ds + \beta \int_0^1 |\mathcal{C}''(s)| ds - \eta \int_0^1 |\nabla f(\mathcal{C}(s))|^2 ds, \tag{4.1}$$

where $f : \Omega \to \mathbb{R}$ denotes a given image and $\mathcal{C}(s) : [0, 1] \to \Omega$ is a parameterized curve and $\alpha, \beta$, and $\eta$ are some positive tuning parameters. The first two terms impose regularity restriction on the contour while the third one denotes the drive induced by the given image. As the image $f$ has large gradient along object boundaries, the functional $E(\mathcal{C})$ will take a small value when the active contour $\mathcal{C}$ resides on these boundaries.

Mumford and Shah [29] proposed minimizing the following functional:

$$\mathbf{E}(u, K) = \int_{\Omega \setminus K} |\nabla u|^2 dx + \eta \int_\Omega (u - f)^2 dx + \mu \text{Length}(K) \tag{4.2}$$

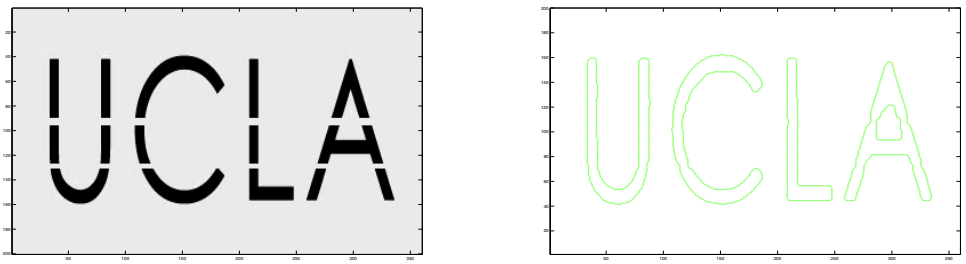
with respect to both the function $u$ defined on $\Omega$ and the boundary $K \subset \Omega$. $\eta, \mu$ are positive tuning parameters.

The segmentation model of Chan-Vese [9] can be expressed as the minimization of the following functional:

$$\mathbf{E}_{CV}(\phi, c_1, c_2) = \int_\Omega \mu(f-c_1)^2 H(\phi) + (f-c_2)^2(1-H(\phi)) + \eta \int_\Omega |\nabla H(\phi)|, \tag{4.3}$$

where $\phi$ is a level set function whose zero level curve presents the segmentation boundary, $H(\cdot)$ is the Heaviside function, $c_1, c_2$ are two scalars, and $\mu, \eta$ are positive parameters. The parameter $\mu$ is often set to be 1 in many applications. If the minimizer of the objective functional in the Mumford-Shah's model is restricted to be $u = c_1 H(\phi) + c_2(1-H(\phi))$, a "binary image", one can easily get Chan-Vese's model.

In Chan-Vese's model, the first two terms are the fitting terms while the third one represents the length of the segmentation boundary. As discussed in [29], the length term prohibits the excessive segmentation boundaries obtained by the Chan-Vese model. Moreover, it also imposes regularity on the boundaries. Chan-Vese model has proven to be an effective segmentation model. However, the length regularization term is insufficient to accomplish the segmentation task under some circumstances. For instance, as shown in Figure 4.1(A), parts of the letters "UCLA" are erased. Even though one can easily recognize the four letters, existing segmentation models, such as Chan-Vese's model, might often capture the existing boundaries instead of restoring the missing ones as illustrated in Figure 4.1(B). In inpainting problems [11], missing image information is also recovered but within given regions assigned in advance. In contrast, we intend to have a segmentation model that can interpolate the missing boundaries automatically without specifying the regions.



(A) (B)

Figure 4.1: Incomplete letters "UCLA" and its integrate segmentation.

To this end, we employ Euler's elastica as a new regularization for the segmentation contour to replace the length term in Chan-Vese's model and get the following functional:

$$\begin{aligned} \mathbf{E}_{CVE}(\phi, c_1, c_2) &= \int_\Omega \mu(f-c_1)^2 H(\phi) + (f-c_2)^2(1-H(\phi)) \\ &\quad + \left[a + b\left(\nabla \cdot \frac{\nabla\phi}{|\nabla\phi|}\right)^2\right] |\nabla H(\phi)|, \end{aligned} \tag{4.4}$$

where $\mu, a, b$ are positive parameters. For $\phi$ being the signed distance level set function, it can be proven that the last term equals to the Euler's elastica energy of the segmentation curve. Specifically, the parameter $\mu$ has a more important role in this modified model than in Chan-Vese's model. It can relax the competition of the fitting term and the Euler's elastica term, aiming to complete missing boundaries as shown in Figure 4.1(B). The parameters $a, b$ control the length and curvature of segmentation boundary. This regularization was originally proposed and used in the famous work of segmentation with depth by Nitzberg, Mumford, and Shiota [28]. It has also been used in the inpainting problem [11] and the illusory contour problem [51, 50]. Recently, in [35], Schoenemann et al. developed a numerical method to minimize the curvature dependent functionals by using linear programming method. In their work, they also considered Chan-Vese's

model with the substitution of Euler's elastica for the length term. In this section, we shall show the details on using the techniques developed in the earlier sections to minimize the Euler's elastic energy for the CVE model (4.4).

In the current work, we use the same technique as in Section 3.3 to deal with the curvature term in the functional (4.4). Note that the functional (4.4) involves the level set function $\phi$, however, only the sign of this function, $H(\phi)$, is needed for the segmentation problem. Following the ideas of the binary level set representation of [23], we introduce a new function $u = H(\phi)$. This was also used in [10] for finding the global minimizer of Chan-Vese's model. More general binary level set representations with global minimization techniques have been developed [45, 47, 44, 2, 46] through some beautiful connections between graph cuts, binary labeling and continuous max-flow problems. As $\nabla \cdot \frac{\nabla H(\phi)}{|\nabla H(\phi)|} = \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|}$, one can rewrite the functional (4.4) to be

$$\mathbf{E}(u, c_1, c_2) = \int_\Omega \mu(f - c_1)^2 u + (f - c_2)^2(1 - u) + \left[a + b\left(\nabla \cdot \frac{\nabla u}{|\nabla u|}\right)^2\right] |\nabla u|, \tag{4.5}$$

where $u$ is supposed to take on either 0 or 1. But note that the curvature makes sense for smooth functions. To fix this issue, as in [2, 10], one can relax the restriction on $u$ to be $0 \leq u \leq 1$. To minimize the functional (4.5), one considers the following constrained minimization problem

$$\begin{aligned} &\min_{u,\mathbf{p},\mathbf{n},c_1,c_2} \int_\Omega \mu(f - c_1)^2 u + (f - c_2)^2(1 - u) + \left[a + b\,(\nabla \cdot \mathbf{n})^2\right] |\mathbf{p}|, \\ &\text{with } \mathbf{p} = \nabla u, |\mathbf{p}| = \mathbf{p} \cdot \mathbf{n}, |\mathbf{n}| \leq 1, u \in [0, 1]. \end{aligned} \tag{4.6}$$

We then construct the following augmented Lagrangian functional:

$$\begin{aligned} \mathcal{L}(v, u, \mathbf{p}, \mathbf{n}, \mathbf{m}, c_1, c_2; \lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_4) &= \int_\Omega \mu(f - c_1)^2 v + (f - c_2)^2(1 - v) + \left[a + b\,(\nabla \cdot \mathbf{n})^2\right] |\mathbf{p}| \\ &+ r_1 \int_\Omega (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) + \int_\Omega \lambda_1 (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) \\ &+ \frac{r_2}{2} \int_\Omega |\mathbf{p} - \nabla u|^2 + \int_\Omega \boldsymbol{\lambda}_2 \cdot (\mathbf{p} - \nabla u) \\ &+ \frac{r_3}{2} \int_\Omega (v - u)^2 + \int_\Omega \lambda_3 (v - u) + \delta_{\mathcal{D}}(v) \\ &+ \frac{r_4}{2} \int_\Omega |\mathbf{n} - \mathbf{m}|^2 + \int_\Omega \boldsymbol{\lambda}_4 \cdot (\mathbf{n} - \mathbf{m}) + \delta_{\mathcal{R}}(\mathbf{m}), \end{aligned} \tag{4.7}$$

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:

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

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

Moreover, $r_i, i = 1, ..., 4$ are positive parameters while $\lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_4$ are Lagrange multipliers. In this augmented Lagrangian functional, as was explained in Section 3.3, the new variable $\mathbf{m}$ is introduced to simplify the associated subproblem on $\mathbf{p}$. As $\mathbf{m}$ is required to be inside $\mathcal{R}$, $|\mathbf{m}| \leq 1$, then $|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m} \geq 0$ for any $\mathbf{p}$, and $|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m} = 0$ if and only if $\mathbf{m} = \frac{\mathbf{p}}{|\mathbf{p}|}$. This avoids the term $\int_\Omega (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m})^2$, which results in a relatively complex functional on $\mathbf{p}$. Moreover, by using the new variable $\mathbf{m}$, the minimizer of the functional related to $\mathbf{p}$ can be obtained exactly and explicitly by using some appropriate shrinkage.

It is well known that some saddle point of the augmented Lagrangian functional (4.7) relates to a minimizer of the functional (4.5). Therefore, one just needs to find the saddle points of the augmented functional. The minimization energy functional for the CVE model given in (4.6) is very similar to the Euler's elastica energy of Section 3.3. We could use an alternating minimization procedure to approximately minimize the variables $u, v, \mathbf{p}, \mathbf{n}, c_1, c_2$ and use a simple gradient ascent method to update the Lagrange multipliers. Algorithms and numerical performance are exposed in [54].

## 5. Tuning of the penalization parameters for ALM

The values of the penalization parameters $r_i$ are very important. They influence the convergence as well as the speed of convergence of the proposed algorithms. Fortunately, there exist some very easy techniques to find the proper intervals for the values of these penalization parameters. Here we review some details.

The ROF model associated with the energy functional (3.1) is convex, thus the ALM is convergent for any positive values of $r_i$ used in Algorithm 1. However, the speed of the convergence depends on the values of $r_i$. Choosing $r_i$ too big or too small could result in more iterations for the solution to converge to the same stopping cretria. For the other higher order models discussed earlier, the values of $r_i$ also influence speed of convergence. In addition, for non-covex energy functionals, some of the penalization parameters need to be sufficiently large to guarantee the convergence.

Fortunately, there are good numerical indicators to use for the determination of the values of $r_i$. This makes it very easy to tune the penalization parameters. The indicators are related to the constraint errors and the decay of the energy functional value. Let us take Algorithm 4 as an example. First, we need to monitor the constraint errors:

$$(R_1^k, R_2^k, R_3^k, R_4^k) = (\|R_1^k\|_{L^1}/\|R_1^0\|_{L^1}, \|R_2^k\|_{L^1}/\|R_2^0\|_{L^1}, \|R_3^k\|_{L^1}/\|R_3^0\|_{L^1}, \|R_4^k\|_{L^1}/\|R_4^0\|_{L^1}), \tag{5.1}$$

with

$$\begin{aligned} R_1^k &= |\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}^k, \\ R_2^k &= \mathbf{p}^k - \langle \nabla u^k, 1 \rangle, \\ R_3^k &= q^k - \partial_x n_1^k - \partial_y n_2^k, \\ R_4^k &= \mathbf{n}^k - \mathbf{m}^k, \end{aligned}$$

Note that all the errors are normalized by scaling the errors with their values from the first iteration. In addition, we also need to monitor the value of the energy functional. Here are some "troubleshooting" tips on how to tune the parameters $r_i$:

- Step 1 Take some reasonable guess for the values of all the $r_i$ and run the algorithms until the stopping criteria are satisfied.
- Step 2 Tune the values of $r_i$ so that the constraint errors $R_i^k$ converge to zero with nearly the same speed asymptotically. If $R_i^k$ goes to zero slower than the others, then increase the value of $r_i$. If $R_i^k$ goes to zero quicker than the others, then decrease the value of $r_i$. It is possible that these constraints errors "behave" rather chaotically in the starting phase of the iterations. However, they shall converge to zero asymptotically with the same "speed" if the values of the $r_i$ are chosen correctly.
- Step 3 By choosing the penalization values $r_i$ sufficiently large, it is always possible to make the constraint errors go to zero. However, the energy functional value may stay large all the time. For ALM, it is not possible to guarantee that the energy functional will decrease monotonically. However, the energy will decrease and then stay at a constant value if $r_i$ are chosen correctly. Thus, if the constraint errors are decreasing correctly, but not the energy functional value, then reduce all the $r_i$ and repeat this tuning process from step 2.

We also need to stop the iterations properly. In all our numerical experiments, we use the relative residuals (5.1), the relative errors of Lagrange multipliers and value of $E(u^k)$ as the stopping criteria. To check the convergence of the iteration process, we first check on $R_i^k$. As in [37], we also check the relative errors of Lagrange multipliers:

$$(L_1^k, L_2^k, L_3^k, L_4^k) = \left( \frac{\|\lambda_1^k - \lambda_1^{k-1}\|_{L^1}}{\|\lambda_1^{k-1}\|_{L^1}}, \frac{\|\boldsymbol{\lambda}_2^k - \boldsymbol{\lambda}_2^{k-1}\|_{L^1}}{\|\boldsymbol{\lambda}_2^{k-1}\|_{L^1}}, \frac{\|\lambda_3^k - \lambda_3^{k-1}\|_{L^1}}{\|\lambda_3^{k-1}\|_{L^1}}, \frac{\|\boldsymbol{\lambda}_4^k - \boldsymbol{\lambda}_4^{k-1}\|_{L^1}}{\|\boldsymbol{\lambda}_4^{k-1}\|_{L^1}} \right), \tag{5.2}$$

and the relative error of the solution $u^k$

$$\frac{\|u^k - u^{k-1}\|_{L^1}}{\|u^{k-1}\|_{L^1}}. \tag{5.3}$$

Besides all these quantities, we also consider how the energy (2.3) is evolving during the iterations by tracking the value of $E(u^k)$. If all the residual errors $R_i^k$ satisfy the stopping criteria $R_i^k < \epsilon_r$ for some given small threshold $\epsilon_r$, the relative errors for the multipliers and the solution $u$ have been reduced to a sufficiently small level (normally can be close to machine accuracy) and the energy functional $E(u^k)$ has come to a steady constant value, then the algorithm has reached a steady state and we can stop the iterations.

## References


[1] Luigi Ambrosio and Simon Masnou. A direct variational approach to a problem arising in image reconstruction. *Interfaces and Free Boundaries*, 5(1):63–82, 2003.

[2] Egil Bae, Jing Yuan, and Xue-Cheng Tai. Global minimization for continuous multiphase partitioning problems using a dual approach. *International Journal of Computer Vision*, pages 1–18, 2010.

[3] Kristian Bredies. Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. *SFB Report*, 6, 2012.

[4] Kristian Bredies, Thomas Pock, and Benedikt Wirth. Convex relaxation of a class of vertex penalizing functionals. *Journal of mathematical imaging and vision*, 47(3):278–302, 2013.

[5] Carlos Brito-Loeza and Ke Chen. On high-order denoising models and fast algorithms for vector-valued images. *Image Processing, IEEE Transactions on*, 19(6):1518–1527, 2010.

[6] Jeff Calder, A Mansouri, and Anthony Yezzi. Image sharpening via sobolev gradient flows. *SIAM Journal on Imaging Sciences*, 3(4):981–1014, 2010.

[7] Antonin Chambolle and Pierre-Louis Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.

[8] RH Chan, A Lanza, S Morigi, and F Sgallari. An adaptive strategy for the restoration of textured images using fractional order regularization. *Numerical Mathematics: Theory, Methods & Applications*, 6(1), 2013.

[9] T. Chan and L.A. Vese. Active contours without edges. *IEEE Trans Image Proc.*, 10:266–277, 2001.

[10] Tony F. Chan, Selim Esedoglu, and Mila Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J. Appl. Math.*, 66(5):1632–1648 (electronic), 2006.

[11] Tony F Chan, Sung Ha Kang, and Jianhong Shen. Euler's elastica and curvature-based inpainting. *SIAM Journal on Applied Mathematics*, pages 564–592, 2002.

[12] Eduardo Cuesta, Mokhtar Kirane, and Salman A Malik. Image structure preserving denoising using generalized fractional time integrals. *Signal Processing*, 92(2):553–563, 2012.

[13] Stephan Didas, Joachim Weickert, and Bernhard Burgeth. Properties of higher order nonlinear diffusion filtering. *Journal of mathematical imaging and vision*, 35(3):208–226, 2009.

[14] Yuping Duan and Weimin Huang. A fixed-point augmented lagrangian method for total variation minimization problems. *Journal of Visual Communication and Image Representation*, 24(7):1168–1181, 2013.

[15] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, volume 9. Society for Industrial Mathematics, 1989.

[16] T. Goldstein and S. Osher. The split bregman method for l1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[17] John B Greer and Andrea L Bertozzi. Traveling wave solutions of fourth order pdes for image processing. *SIAM Journal on Mathematical Analysis*, 36(1):38–68, 2004.

[18] Patrick Guidotti and Kate Longo. Two enhanced fourth order diffusion models for image denoising. *Journal of Mathematical Imaging and Vision*, 40(2):188–198, 2011.

[19] Langhua Hu, Duan Chen, and Guo-Wei Wei. High-order fractional partial differential equation transform for molecular surface construction. *Molecular based mathematical biology*, 1:1–25, 2013.

[20] P Jidesh and Santhosh George. Fourth–order variational model with local–constraints for denoising images with textures. *International Journal of Computational Vision and Robotics*, 2(4):330–340, 2011.

[21] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[22] Ron Kimmel, Ravi Malladi, and Nir Sochen. Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images. *International Journal of Computer Vision*, 39(2):111–129, 2000.

[23] Johan Lie, Marius Lysaker, and Xue Tai. A binary level set model and some applications to mumford-shah image segmentation. *IEEE Transactions on Image Processing*, 15(5):1171–1181, 2006.

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

[25] Kent Andre Mardal, Xue-Cheng Tai, and Ragnar Winther. A robust finite element method for darcy-stokes flow. *SIAM Journal on Numerical Analysis*, pages 1605–1631, 2003.

[26] Lihua Min, Xiaoping Yang, and Changfeng Gui. Entropy estimates and large-time behavior of solutions to a fourth-order nonlinear degenerate equation. *Communications in Contemporary Mathematics*, 15(04), 2013.

[27] Lihua Min, Xiaoping Yang, and Dong Ye. Well-posedness for a fourth order nonlinear equation related to image processing. *Nonlinear Analysis: Real World Applications*, 2013.

[28] D Mumford, M Nitzberg, and T Shiota. Filtering, segmentation and depth. *Lecture Notes in Computer Science*, 662, 1993.

[29] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math, 42*, 42:577–685, 1989.
[30] Ehsan Nadernejad and Søren Forchhammer. Wavelet-based image enhancement using fourth order pde. In *Intelligent Signal Processing (WISP), 2011 IEEE 7th International Symposium on*, pages 1–6. IEEE, 2011.
[31] Carola-Bibiane Papafitsorosand Schönlieb. A combined first and second order variational approach for image reconstruction. *Journal of Mathematical Imaging and Vision*, 48(2):308–338, 2014.
[32] Guy Rosman, Alex M Bronstein, Michael M Bronstein, Xue-Cheng Tai, and Ron Kimmel. Group-valued regularization for analysis of articulated motion. In *Computer Vision–ECCV 2012. Workshops and Demonstrations*, pages 52–62. Springer, 2012.
[33] Guy Rosman, Yu Wang, Xue-Cheng Tai, Ron Kimmel, and Alfred M Bruckstein. Fast regularization of matrix-valued images. In *Computer Vision–ECCV 2012*, pages 173–186. Springer, 2012.
[34] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
[35] T. Schoenemann, F. Kahl, and D. Cremers. Curvature regularity for region-based image segmentation and inpainting: A linear programming relaxation. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 17–23. IEEE, 2009.
[36] Carola-bibiane Schönlieb and Andrea Bertozzi. Unconditionally stable schemes for higher order inpainting. *Communications in Mathematical Sciences*, 9(2):413–457, 2011.
[37] X.C. Tai, J. Hahn, and G.J. Chung. A fast algorithm for euler's elastica model using augmented lagrangian method. *SIAM Journal on Imaging Sciences*, 4:313, 2011.
[38] Xue-Cheng Tai and Chunlin Wu. Augmented lagrangian method, dual methods and split bregman iteration for rof model. *Scale Space and Variational Methods in Computer Vision*, pages 502–513, 2009.
[39] Wekipedia. Co-area formula (http://en.wikipedia.org/wiki/coarea_formula), 2013.
[40] Wikipedia. Total variation (http://en.wikipedia.org/wiki/total_variation), 2014.
[41] C. Wu and X.C. Tai. Augmented lagrangian method, dual methods, and split bregman iteration for rof, vectorial tv, and high order models. *SIAM Journal on Imaging Sciences*, 3(3):300–339, 2010.
[42] Fenlin Yang, Ke Chen, and Bo Yu. Efficient homotopy solution and a convex combination of rof and llt models for image restoration. *International Journal of Numerical Analysis & Modeling*, 9(4), 2012.
[43] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for \ell_1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008.
[44] J. Yuan, E. Bae, and X.C. Tai. A study on continuous max-flow and min-cut approaches. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2217–2224. IEEE, 2010.
[45] J. Yuan, E. Bae, X.C. Tai, and Y. Boykov. A continuous max-flow approach to potts model. *Computer Vision–ECCV 2010*, pages 379–392, 2010.
[46] Jing Yuan, Egil Bae, Xue-Cheng Tai, and Yuri Boykov. A spatially continuous max-flow and min-cut framework for binary labeling problems. *Numerische Mathematik*, pages 1–29, 2013.
[47] Jing Yuan, Juan Shi, and Xue-Cheng Tai. A convex and exact approach to discrete constrained tv-l1 image approximation. *East Asian Journal on Applied Mathematics*, to appear.
[48] Weili Zeng, Xiaobo Lu, and Xianghua Tan. Non-linear fourth-order telegraph-diffusion equation for noise removal. *IET Image Processing*, 7(4):335–342, 2013.
[49] W. Zhu and T. Chan. Image denoising using mean curvature of image surface. *SIAM Journal on Imaging Sciences*, 5(1):1–32, 2012.
[50] W. Zhu, T. Chan, and Selim Esedoglu. Segmentation with depth: A level set approach. *SIAM Journal on Scientific Computing*, 28(5):1957–1973, 2006.
[51] Wei Zhu and Tony Chan. A variational model for capturing illusory contours using curvature. *Journal of Mathematical Imaging and Vision*, 27(1):29–40, 2007.
[52] Wei Zhu, Xue-Cheng Tai, and Tony Chan. Augmented lagrangian method for a mean curvature based image denoising model. *Inverse Problems and Imaging*, 7(3):1075–1097, 2012.
[53] Wei Zhu, Xue-Cheng Tai, and Tony Chan. Augmented lagrangian method for a mean curvature based image denoising model. *Inverse Problems and Imaging*, 7(4):1409–1432, 2013.
[54] Wei Zhu, Xue-Cheng Tai, and Tony Chan. Image segmentation using euler's elastica as the regularization. *Journal of Scientific Computing*, 57(2):414–438, 2013.

Department of Mathematics , University of Bergen, Bergen, Norway • Tai@math.uib.no