On Variable Splitting and Augmented Lagrangian Method for Total Variation Related Image Restoration Models

Zhifang Liu, Yuping Duan, Chunlin Wu and Xue-Cheng Tai

Introduction

This short survey provides a brief review of the variable splitting and augmented Lagrangian method for total variation (TV) related image restoration models. We will focus on this computational problem closely, and do not plan to touch other related topics like theoretical model analysis and algorithmic connections, which can be referred to, e.g., Aubert and Kornprobst (2010); Glowinski et al (2016a) and references therein. Also, to keep the context as compact as possible, we would not expand all the details, although there are definitely lots of exellent works in the literature.

Total variation, which is a semi-norm of the space of functions of bounded variation, was first proposed for image denoising by Rudin, Osher and Fatemi (ROF) in Rudin et al (1992). In the discrete setting, it is essentially the L_1 norm of gradients and can maintain the sparse discontinuities. Therefore, it is appropriate to preserve image edges Strong and Chan (2003); Chambolle and Lions (1997); Caselles et al (2007) that are usually the most important features for images to recover. Owing to its edge-preserving property and convexity, total variation has been demonstrated very successful and become popular in image restoration like image denois-

Zhifang Liu
School of Mathematical Sciences, Tianjin Normal University, Tianjin, China
e-mail: matlzhf@tjnu.edu.cn

Yuping Duan
Center for Applied Mathematics, Tianjin University, Tianjin, China
e-mail: yuping.duan@tju.edu.cn

Chunlin Wu (✉ Corresponding author)
School of Mathematical Sciences, Nankai University, Tianjin, China
e-mail: wucl@nankai.edu.cn

Xue-Cheng Tai
Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
e-mail: xuechengtai@hkbu.edu.hk

ing Rudin et al (1992, 2003); Aubert and Aujol (2008); Le et al (2007); Jin and Yang (2011); Dong and Zeng (2013); Sciacchitano et al (2015); Zhao et al (2014), image deblurring Rudin and Osher (1994); Chan and Wong (1998); Chan et al (2005); Huang et al (2008); Ma et al (2013) and image inpainting Shen and Chan (2002); Chan et al (2005); Bertalmio et al (2003), and also various other types of image processing tasks including image decomposition Aujol and Chambolle (2005); Aujol et al (2005); Le and Vese (2005); Vese and Osher (2004); Ng et al (2013); Vese and Osher (2003), image segmentation Chan and Vese (2001); Chan et al (2006), image super-resolution Marquina and Osher (2008), face recognition Chen et al (2006a), fluorescence tomography Freiberger et al (2010), CT reconstruction Persson et al (2001); Jia et al (2011); Tian et al (2011); Sidky and Pan (2008); Sidky et al (2011); Ritschl et al (2011); Zhu et al (2012); Yang et al (2010); Chen et al (2013), phase retrieval Chang et al (2016, 2018) and so on.

The total variation model has been generalized in many ways for different purposes. The original total variation regularization was proposed for gray image restoration Rudin et al (1992), which is the single channel case. To restore multichannel data, such as color images with RGB channels, people extended it to color TV and vectorial TV regularizations Blomgren and Chan (1998); Sapiro and Ringach (1996); Bresson and Chan (2008). It is well known that images recovered by total variation regularized models have the undesired staircase effect. To prevent the total variation oversharpening, there are several remarkable methods to improve the total variation regularization. These include the variable exponent TV models Chen et al (2006b); Li et al (2010) and a wide class of high order models, such as inf-convolution model Chambolle and Lions (1997), Chan-Marquina-Mulet model Chan et al (2000), Lysaker-Lundervold-Tai model Lysaker et al (2003) and total generalized variation model Bredies et al (2010); Bredies and Holler (2020), etc. By co-area formula, the total variation is the intrgal of lengths of all level curves of the intensity function Chan and Shen (2005). One natural extension way is thus to introduce curve curvature term for regularization. For example, Euler's elastica which contains both lengths and curvatures was proposed for image inpainting Masnou and Morel (1998); Chan et al (2002); Tai et al (2011); Yashtini and Kang (2016), denoising Tai et al (2011); Duan et al (2013); Zhang et al (2017b), segmentation Esedoglu and March (2003); Zhu et al (2006, 2013b); Duan et al (2014); Bae et al (2017); Tan et al (2020), zooming Tai et al (2011); Duan et al (2013), illusory contour Masnou and Morel (2005); Kang et al (2014), image decomposition Liu et al (2018) and image reconstruction Zhang et al (2017a); Yan and Duan (2020). Such regularity can provide strong priors for the continuity of edges. Another total variation related geometric regularization technique we would like to mention is mean curvature minimization Zhu and Chan (2012), which considers the image or graph in a high-dimensional space and transfers the image minimization problems to the corresponding surface minimization problems. From the viewpoint of image domain, total variation regularization was also extended to implicit surfaces, triangulated meshes and even general manifolds for image and data processing on curved spaces Bertalmio et al (2001); Lai and Chan (2011); Wu et al (2012); Lellmann et al (2013); Wu et al (2013); Weinmann et al (2014); Osher et al (2017) and nor-

mal vector filtering for surface denoising Zhang et al (2015); Liu et al (2019). By exploiting the spatial interactions in images, total variation regularization was also generalized to nonlocal TV Peyré et al (2008); Lou et al (2010); Dong et al (2012); Liu and Zheng (2017) and total fractional-order variation Zhang et al (2015). By using non-convex penalty functions instead of the L_1 norm, non-convex TV regularizations got more and more attentions in recent years, see Blake and Zisserman (1987); Nikolova (2005); Nikolova et al (2008, 2010); Xu et al (2011); Chen et al (2012b); Hintermüller and Wu (2013); Coll et al (2015); Bian and Chen (2015); Lanza et al (2016); Chan et al (2018); Feng et al (2018); Wu et al (2018); Zeng and Wu (2018, 2019); Zeng et al (2019a,b); You et al (2019); Xu et al (2019); Selesnick et al (2020); Pang et al (2020); Gao and Wu (2020) and the references therein. They have been shown capable to generate good results with neat edges, as indicated by the interesting the lower bound theory Nikolova (2005); Chen et al (2010, 2012b); Feng et al (2018); Zeng and Wu (2018, 2019); Zeng et al (2019b).

However, the non-smoothness of the total variation semi-norm gives rise to a challenge of its minimization. To overcome this problem, the common way is replacing total variation by its smoothed versions in image restoration model. Therefore, one can solve the new associated Euler-Lagrangian equation and obtain an approximate solution of the original model Acar and Vogel (1994); Blomgren et al (1997). For solving this Euler-Lagrangian equation, Rudin, Osher and Fatemi proposed a gradient flow method Rudin et al (1992). This method is slow due to strict constraints on the time step size and many methods have been proposed to improve on it. Some efficient methods are fixed point methods Vogel and Oman (1998); Chan and Mulet (1999), dual methods Chan et al (1999); Chambolle (2004); Dong et al (2009); Esser et al (2010); Chambolle and Pock (2011); Cai et al (2013); He et al (2014); Wen et al (2016), the split Bregman method Goldstein and Osher (2009) and splitting-and-penalty based methods Wang et al (2008); Yang et al (2009b,a); Guo et al (2009), proximity algorithms Micchelli et al (2011, 2013); Chen et al (2014), alternating direction method of multipliers Chan et al (2013); Ng et al (2010); Wahlberg et al (2012); Ng et al (2013) and augmented Lagrangian methods Tai and Wu (2009); Wu and Tai (2010); Wu et al (2011); Zhang and Wu (2011); Wu et al (2012).

The augmented Lagrangian method was originally introduced by Hestenes (1969) and Powell (1969) for solving constrained optimization problem and further systematically studied by many researchers, such as Rockafellar Rockafellar (1974, 1976) and Bertsekas Bertsekas (1996(first published 1982). It was also widely applied to optimize unconstrained minimization problem with the aid of operator splitting technique Glowinski (2015) by which one can transform the unconstrained optimization problem to its equivalent constrained versions. One of the special and very useful instance of augmented Lagrangian methods is the alternating direction method of multipliers (ADMM)Boyd (2010); He and Yuan (2012); Glowinski (2014, 2015); Glowinski et al (2016b); Yan and Yin (2016), which is famous in optimization and statistics community and has broad applications. ADMM has been extensively studied in recent decades He et al (2002); Boyd (2010); He and Yuan (2012, 2015); He et al (2016); Yang and Han (2016); Jiao et al (2016); Cui et al

(2016); Yue et al (2018); He et al (2020) and has many practical variants, such as linearized ADMM Wang and Yuan (2012); He et al (2020), preconditioned ADMM Esser et al (2010); Deng and Yin (2016); Jiao et al (2017), proximal ADMM Fazel et al (2013); Chen et al (2015); Li et al (2016), accelerated ADMM Ouyang et al (2015); Kadkhodaie et al (2015); Li and Lin (2019), stochastic ADMM Fang et al (2017); Chen et al (2018); Ouyang et al (2013) and non-convex ADMM Chartrand and Wohlberg (2013); Yang et al (2017); Li and Pong (2015); Houska et al (2016); Hong et al (2016); Wang et al (2018, 2019); Themelis and Patrinos (2020); Boţ and Nguyen (2020).

Indeed, the variable splitting and augmented Lagrangian method gained great successes in solving nonlinear variational problems that arise from physics, mechanics, economics, etc Glowinski and Tallec (1989); Glowinski (1984). The variable splitting step helps to transform a complicated problem into a constrained optimization with more variables, then an iteration based on augmented Lagrangian method is performed with several easier subproblems. Inspired by this, the method was proposed by Tai and Wu to optimize the total variation based image restoration model in Tai and Wu (2009); Wu and Tai (2010). As expected, augmented Lagrangian methods benefit from the periodic boundary condition which is commonly assumed for image processing problems and the L_1 norm which is included in the total variation semi-norm. The augmented Lagrangian method for TV based image restoration model has two subproblems. The periodic boundary condition allows us to solve one of the subproblems via Fourier transformation with FFT implementation in the case of deconvolution case. Meanwhile, the other subproblem with the L^1 norm has closed form solution. Despite of the fact that the image processing problems are naturally in large scale, these two advantages of the augmented Lagrangian method make it efficient in minimizing the objective functionals related with the non-smooth total variation for various image processing tasks. Since Tai and Wu (2009); Wu and Tai (2010), the variable splitting and augmented Lagrangian method has been widely applied to total variation related minimizations like the single channel case Wu and Tai (2010); Tai and Wu (2009); Wu et al (2011), the multichannel case Wu and Tai (2010); Zhang and Wu (2011), high order models Wu and Tai (2010); Gao et al (2018), TV-Stokes model Hahn et al (2012); Chen et al (2012a), wavelet based image restoration Afonso et al (2010, 2011); Dong and Zhang (2013), Euler's elastica image restoration model Tai et al (2011); Duan et al (2013); Zhang et al (2017b); Yashtini and Kang (2016), Euler's elastica image segmentation model Zhu et al (2013b); Bae et al (2017), mean curvature image denoising Zhu et al (2013a); Sun and Chen (2014); Myllykoski et al (2015); Zhang (2018), total variation video restoration Chan et al (2011), MR image reconstruction Guo and Huang (2009); Ramani and Fessler (2011); Allison et al (2013), SAR imaging Güven et al (2016), total variation minimization in curved spaces for either data processing Bertalmio et al (2001); Lai and Chan (2011); Wu et al (2012); Lellmann et al (2013); Wu et al (2013); Weinmann et al (2014); Osher et al (2017) or normal-vector-filtering based surface denoising Zhang et al (2015); Liu et al (2019) and even more in Freiberger et al (2010); Ilbey et al (2017); Koko and Jehan-Besson (2010); Han et al (2014); Li et al (2013); Freiberger et al (2010); Nien and Fessler (2015). Therein for some

complicated non-convex models like Euler's elastica or mean curvature based, how to introduce the auxiliary variables is tricky and important to get stable and efficient algorithms. There are some close connections between the augmented Lagrangian method and other approaches such as split Bregman method Goldstein and Osher (2009) and Chambolle's projection method Chambolle (2004), and some works for improving classical augmented Lagrangian method can be found in Li et al (2013); Xiao and Song (2012), etc.

The content included here are organized as follows. In section 2, we present some basic notations. In section 3, we present augmented Lagrangian methods TV restoration models with L^2 fidelity term and TV restoration models with non-quadratic fidelity. In section 4, we present augmented Lagrangian methods for multichannel TV restoration. In section 5, we present augmented Lagrangian methods for high order models, including Lysaker-Lundervold-Tai model, total generalized variation model, Euler's elastica model and mean curvature model. In section 6, we show some numerical experiments. We conclude this paper in section 7.

Basic notation

We follow Wu and Tai (2010) for most notations. As a gray image is a 2D array, we represent it by an $N \times N$ matrix, without the loss of generality. It is useful to denote the Euclidean space $\mathbb{R}^{N\times N}$ as $\mathscr{X}$, and write $\mathscr{Y} = \mathscr{X} \times \mathscr{X}$. We recall the discrete gradient operator

$$\begin{aligned} \nabla : \mathscr{X} &\to \mathscr{Y} \\ x &\to \nabla x, \end{aligned}$$

where ∇x is given by

$$(\nabla x)_{i,j} = ((\mathring{D}_1^+ x)_{i,j}, (\mathring{D}_2^+ x)_{i,j}), i, j = 1, \ldots, N,$$

with

$$(\mathring{D}_1^+ x)_{i,j} = \begin{cases} x_{i,j+1} - x_{i,j}, & 1 \le j \le N-1, \\ x_{i,1} - x_{i,N}, & j = N, \end{cases}$$

$$(\mathring{D}_2^+ x)_{i,j} = \begin{cases} x_{i+1,j} - x_{i,j}, & 1 \le i \le N-1, \\ x_{1,j} - x_{N,j}, & i = N. \end{cases}$$

Here $\mathring{D}_1^+$ and $\mathring{D}_2^+$ are used to denote forward difference operators with periodic boundary condition for FFT algorithm implementation. We mention that other boundary conditions with corresponding implementation tricks can also be adopted.

The usual inner products and L^2 norms in the spaces $\mathscr{X}$ and $\mathscr{Y}$ are as follows. We denote

$$\langle x, z \rangle = \sum_{1 \le i,j \le N} x_{ij} z_{ij} \text{ and } \|x\| = \sqrt{\langle x, x \rangle},$$

for $x, z \in \mathscr{X}$; and

$$\langle w, y \rangle = \langle w^1, y^1 \rangle + \langle w^2, y^2 \rangle, \text{ and } \|y\| = \sqrt{\langle y, y \rangle},$$

for $y = (y^1, y^2) \in \mathscr{Y}$ and $w = (w^1, w^2) \in \mathscr{Y}$. At each pixel (i, j), we define

$$|y_{i,j}| = |(y^1_{i,j}, y^2_{i,j})| = \sqrt{(y^1_{i,j})^2 + (y^2_{i,j})^2}$$

as the usual Euclidean norm in $\mathbb{R}^2$. We mention that $\|x\|_{L^p}$ is used to denote the general L^p norm of $x \in \mathscr{X}$.

By using the inner products of $\mathscr{X}$ and $\mathscr{Y}$, it is clear that the discrete divergence operator, as the adjoint operator of $-\nabla$, is as follows

$$\begin{aligned} \operatorname{div} &: \mathscr{Y} \to \mathscr{X} \\ y = (y^1, y^2) &\to \operatorname{div} y, \end{aligned}$$

where

$$(\operatorname{div} y)_{i,j} = y^1_{i,j} - y^1_{i,j-1} + y^2_{i,j} - y^2_{i-1,j} = (\mathring{D}_1^- y^1)_{i,j} + (\mathring{D}_2^- y^2)_{i,j},$$

with backward difference operators $\mathring{D}_1^-$ and $\mathring{D}_2^-$ and periodic boundary conditions $y^1_{i,0} = y^1_{i,N}$ and $y^2_{0,j} = y^2_{N,j}$.

Augmented Lagrangian method for total variation related image restoration models

We assume $d \in \mathscr{X}$ to be an observed image. As usual, we model the degradation procedure as

$$\underline{x} \xrightarrow{\text{linear transformation}} K\underline{x} \xrightarrow{\text{noise}} d, \tag{1}$$

where $\underline{x} \in \mathscr{X}$ is the ground truth image and $K : \mathscr{X} \to \mathscr{X}$ is a linear operator like a blur. In other cases, such as when K is a Radon transform or a subsampling, the dimensions of the observed data d and the ground truth data $\underline{x}$ may be different. However, there is no essential difficulty and the method framework here also applies. Here the noise is not necessarily to be additive and could be Gaussian, impulsive, Poisson or even others. The task of image restoration is to recover $\underline{x}$ from d. In this survey we only consider the case where the linear operator K is given. Even so, we usually cannot directly solve $\underline{x}$ from (1), because this is a typical inverse problem. Both the random measurement noise and the bad condition number of K bring computational difficulties. Regularization on the solution should be considered to overcome the ill-posedness.

Although the classical Tikhonov regularization has achieved great successes in lots of general inverse problems, it turns out to over smooth image edges, the most

important image structure. Indeed, one of the most basic and successful image restoration models is based on total variation regularization, which reads

$$\min_{x\in\mathscr{X}}\{E(x)=F(Kx)+R(\nabla x)+B(x)\}, \tag{2}$$

where $F(Kx)$ is a fidelity term, $R(\nabla x)$ is the total variation of x Rudin et al (1992) defined by

$$R(\nabla x)=\mathrm{TV}(x)=\sum_{1\le i,j\le N}|(\nabla x)_{i,j}|, \tag{3}$$

and $B(x)$ is an indicator function of box constraints defined as follows

$$B(x)=\begin{cases}0, & \underline{b}\le x_{i,j}\le \overline{b}, \forall\, i,j,\\ +\infty, & \text{otherwise.}\end{cases}$$

Lots of researches Le et al (2007); Ng et al (2010); Sidky and Pan (2008); Sidky et al (2011); Chan and Ma (2012); Chan et al (2013) show that to involve this kind of constraints is useful, when the intensity range is clear. Otherwise, one can just let the box parameters $\underline{b}$ be $-\infty$ or $\overline{b}$ be $+\infty$. This model includes numerous particular cases studied in the literatures.

For further analysis and interpretation, we make the following assumptions:

- Assumption 1. $\mathrm{Null}(\nabla)\cap\mathrm{Null}(K)=\{0\}$;
- Assumption 2. $\mathrm{dom}(R\circ\nabla)\cap\mathrm{dom}(F\circ K)\cap\mathrm{dom}(B)\neq\varnothing$;
- Assumption 3. $F(z)$ is convex, proper, coercive, and lower semi-continuous;
- Assumption 4. $\mathrm{dom}(F)$ is open,

where $\mathrm{Null}(\cdot)$ is the null space of $\cdot$; $\mathrm{dom}(F)=\{z\in\mathscr{X}:F(z)<+\infty\}$ is the domain of F, and $\mathrm{dom}(R\circ\nabla),\mathrm{dom}(B),\mathrm{dom}(F\circ K)$ are similar. Here we have some comments on these assumptions, which are relatively quite natural. Since most linear operators Ks like blur kernels correspond essentially to averaging operations, Assumption 1 is reasonable. Moreover, although the fidelity terms $F(\cdot)$s are diverse by the statistics of the noise models, many of them meet all of those Assumption 3 and 4, like the following typical ones:

1. The squared L^2 fidelity (corresponding to Gaussian noise):

$$F(Kx)=\frac{\alpha}{2}\|Kx-d\|^2,$$

2. The L^1 fidelity Nikolova (2002, 2004) (corresponding to impulsive noise):

$$F(Kx)=\alpha\|Kx-d\|_{L^1},$$

3. The Kullback-Leibler (KL) divergence fidelity (corresponding to Poisson noise, assuming $d_{i,j}>0,\forall i,j$, as in Le et al (2007)):

$$F(Kx)=\begin{cases}\alpha\sum\limits_{1\le i,j\le N}((Kx)_{i,j}-d_{i,j}\log(Kx)_{i,j}), & (Kx)_{i,j}>0,\forall\, i,j,\\ +\infty, & \text{otherwise,}\end{cases}$$

where $\alpha > 0$ is a parameter. Note for Poisson noise, we use the definition of the fidelity on the whole space $\mathscr{D}$ for analysis convenience, compared to Le et al (2007) (where $K = I$) and Brune et al (2009).

Under the Assumptions 1, 2, 3 and 4, it is not difficult to see that the functional $E(x)$ in (2) is convex, proper, coercive, and lower semi continuous. Thus we have the following existence and uniqueness result, by the generalized Weierstrass theorem and Fermat's rule Glowinski and Tallec (1989); Rockafellar and Wets (1998); Ekeland and Témam (1999).

Theorem 1. *The minimization problem* (2) *has at least one solution x, which satisfies*

$$0 \in K^* \partial F(Kx) - \operatorname{div} \partial R(\nabla x) + \partial B(x), \tag{4}$$

with $\partial F(Kx)$ and $\partial R(\nabla x)$ being the sub-differentials Rockafellar and Wets (1998); Ekeland and Témam (1999) of F at Kx and R at ∇x, respectively. Moreover, if $F \circ K(x)$ is strictly convex, the minimizer is unique.

Next, we present to use the augmented Lagrangian method for TV regularization based image restoration models (2) which satisfy our assumptions.

Augmented Lagrangian method for TV-L^2 *restoration*

In this section, we review the augmented Lagrangian method proposed for the TV restoration model with L^2 fidelity term Tai and Wu (2009); Wu and Tai (2010),

$$\min_{x \in \mathscr{X}} \left\{ E_{\mathrm{TV}}(x) = \frac{\alpha}{2} \|Kx - d\|^2 + R(\nabla x) \right\}, \tag{5}$$

where $\alpha > 0$ and $R(\nabla x)$ is defined as in (3). This model is a special case of model (2), where $F(Kx) = \frac{\alpha}{2}\|Kx - d\|^2$ and the box constraint vanishes. In the literatures, people commonly calls model (5) as TV-L^2 model.

The TV-L^2 model is a fundamental model in image restoration, which is usually applied for removing Gaussian type noise and the linear degradation like blur in image restoration problems Rudin et al (1992); Rudin and Osher (1994); Acar and Vogel (1994). By standard Bayesian estimation, the L^2 fidelity term is deduced from the statistical distribution of the i.i.d Gaussian noise, which guarantees that the recovered image resembles the underly truth image closely. Meanwhile, the total variation regularization preserves the sharp edges.

As we mention before, the total variation term is non-smooth and is a compound of the L^1 norm and the gradient operator. There is a basic idea that is decoupling the total variation term and treating the L^1 norm and the gradient operator relatively independently. By combining with variable splitting technique, the augmented Lagrangian method demonstrates this idea nicely.

First, we introduce an auxiliary variable $y \in \mathscr{Y}$ for ∇x and convert the minimization problem (5) to an equivalent constrained optimization problem

$$\begin{aligned} &\min_{x\in\mathscr{X},y\in\mathscr{Y}} \left\{ G_{\mathrm{TV}}(x,y) = \frac{\alpha}{2}\|Kx-d\|^2 + R(y) \right\}, \\ &\text{s.t.} \quad y = \nabla x. \end{aligned} \tag{6}$$

Then, we define the following augmented Lagrangian function for the constrained optimization problem (6),

$$\mathscr{L}_{\mathrm{TV}}(x,y;\lambda) = \frac{\alpha}{2}\|Kx-d\|^2 + R(y) + \langle \lambda, y-\nabla x\rangle + \frac{\beta}{2}\|y-\nabla x\|^2, \tag{7}$$

with the Lagrange multiplier $\lambda \in \mathscr{Y}$ and a positive penalty parameter β. The augmented Lagrangian method for the problem (6) is to seek a saddle-point of the augmented Lagrangian function (7):

$$\begin{aligned} &\text{Find } (x^*,y^*,\lambda^*) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{Y}, \\ &\text{s.t.} \quad \mathscr{L}_{\mathrm{TV}}(x^*,y^*;\lambda) \le \mathscr{L}_{\mathrm{TV}}(x^*,y^*;\lambda^*) \le \mathscr{L}_{\mathrm{TV}}(x,y;\lambda^*), \\ &\qquad \forall (x,y,\lambda) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{Y}, \end{aligned} \tag{8}$$

The following theorem Glowinski and Tallec (1989); Wu and Tai (2010) reveals the relation between the solution of problem (5) and the saddle-point of problem (8).

Theorem 2. *$x^* \in \mathscr{X}$ is a solution of problem* (5) *if and only if there exist $y^* \in \mathscr{Y}$ and $\lambda^* \in \mathscr{Y}$ such that $(x^*,y^*;\lambda^*)$ is a saddle-point of problem* (8)*.*

Finally, we employ an alternating direction iterative procedure in the augmented Lagrangian method to seek a saddle-point of problem (8); See Algorithm 0.1.

Algorithm 0.1: Augmented Lagrangian method for TV-L^2 model

Initialization: $x^{-1}=0$, $y^{-1}=0$, $\lambda^0=0$;
Iteration: For $k=0,1,\ldots$:

1. compute (x^k,y^k) as an (approximate) minimizer of the augmented Lagrangian function (7) with the Lagrange multiplier λ^k, i.e.,

$$(x^k,y^k) \approx \arg\min_{(x,y)\in\mathscr{X}\times\mathscr{Y}} \mathscr{L}_{\mathrm{TV}}(x,y;\lambda^k), \tag{9}$$

   where $\mathscr{L}_{\mathrm{TV}}(x,y;\lambda^k)$ is defined as (7);
2. update

$$\lambda^{k+1} = \lambda^k + \beta(y^k - \nabla x^k).$$

We can see that the minimization problem (9) still can not be solved directly and exactly. Our strategy is separating the problem (9) into two subproblems with respect to x and y and minimizing them alternatively.

The solution to sub-minimization problem w.r.t. x

Given y, the minimization problem (9) with respect to x is

$$\min_{x\in\mathscr{X}}\left\{\frac{\alpha}{2}\|Kx-d\|^2-\langle\lambda^k,\nabla x\rangle+\frac{\beta}{2}\|y-\nabla x\|^2\right\}.$$

It is a quadratic optimization problem, whose first-order optimization condition gives a linear equation

$$(\alpha K^*K-\beta\Delta)x=\alpha K^*d-\mathrm{div}(\lambda^k+\beta y). \tag{10}$$

If K is a convolution operator like a convolution blur, the above equation under periodic boundary condition can be efficiently solved via Fourier transform with fast Fourier transform (FFT) implementation Wang et al (2008); Wu and Tai (2010). One can obtain its solution by

$$x=\mathscr{F}^{-1}\left(\frac{\alpha\mathscr{F}(K^*)\mathscr{F}(d)-\mathscr{F}(\mathrm{div})\mathscr{F}(\lambda^k+\beta y)}{\alpha\mathscr{F}(K^*)\mathscr{F}(K)-\beta\mathscr{F}(\Delta)}\right),$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ denote the Fourier transform and the inverse Fourier transform. Fourier transforms of operators K^*, K, div, and Δ mean the transforms of the corresponding convolution kernels. If K is not a convolution operator, such as a Radon transform or a subsampling, we can solve the above equation (10) by other well-developed linear solvers like conjugate gradient (CG) method.

The solution to sub-minimization problem w.r.t. y

Given x, the minimization problem (9) with respect to y is

$$\min_{y\in\mathscr{Y}}\left\{R(y)+\langle\lambda^k,y\rangle+\frac{\beta}{2}\|y-\nabla x\|^2\right\}. \tag{11}$$

According to the definition of $R(y)$, we can rewrite (11) as

$$\min_{y\in\mathscr{Y}}\left\{\sum_{1\le i,j\le N}|y_{i,j}|+\frac{\beta}{2}\sum_{1\le i,j\le N}\left|y_{i,j}-\left(\nabla x-\frac{\lambda^k}{\beta}\right)_{i,j}\right|^2\right\}, \tag{12}$$

whose solution is in closed form as follows

$$y_{i,j}=\max\left(0,1-\frac{1}{\beta|\eta_{i,j}|}\right)\eta_{ij}, \tag{13}$$

where $\eta=\nabla x-\lambda^k/\beta\in\mathscr{Y}$. This solution can be derived from the first-order optimization condition via the subdifferential theory Wang et al (2008) or the geometric

explanation of the minimizer Wu et al (2011). We remark that the geometric method can be easily extended to higher (>2) dimensional case Wu and Tai (2010); Wu et al (2011) (See, e.g., multichannel image restoration and high order models in later sections) or the case where $R(\cdot)$ is non-convex Wu et al (2018).

Here, we review the geometric interpretation of the formula (13) given in Wu et al (2011). As one can see, the problem (12) is separable and at each pixel (i, j), we can reduce it to a simple form

$$\min_{u\in\mathbb{R}^2}\left\{|u|+\frac{\beta}{2}|u-v|^2\right\}, \tag{14}$$

where $v\in\mathbb{R}^2$; See Figure 1.

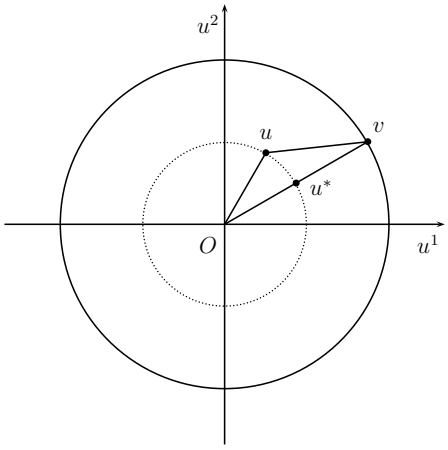


Fig. 1: A geometric interpretation of the formula (13)

In fact, the minimizer of (14) locates in the same quadrant of v and inside of the solid circle with O as center and $|v|$ as radius; See Figure 1. Without loss of generality, we consider the points inside the solid circle at the first quadrant, e.g., u. We draw a dotted circle with O as center and $|u|$ as radius, which intersects the line segment Ov at a point u^*. By the triangle inequality, we have

$$|u|+|u-v|\geq|v|=|u^*|+|u^*-v|.$$

Since $|u|=|u^*|$, we obtain

$$|u-v|\geq|u^*-v|,$$

which indicates

$$|u|+\frac{\beta}{2}|u-v|^2\geq|u^*|+\frac{\beta}{2}|u^*-v|^2.$$

The above equality implies that the solution of (14) locates on the line segment Ov. Therefore, we let $u = \gamma v$ with $0 \leq \gamma \leq 1$ and simplify the problem (14) into an univariate optimization problem

$$\min_{0\leq\gamma\leq 1} \left\{ \gamma|v| + \frac{\beta}{2}(\gamma-1)^2|v|^2 \right\}. \tag{15}$$

The above problem (15) can be solved exactly and has a closed form solution,

$$\gamma^* = \max\left(0, 1 - \frac{1}{\beta|v|}\right).$$

According to (10) and (13), we can solve (9) by an alternating minimization procedure; See Algorithm 0.2.

Algorithm 0.2: Augmented Lagrangian method for TV-L^2 model – solve the minimization problem (9)

Initialization: $x^{k,0} = x^{k-1}$, $y^{k,0} = y^{k-1}$;
Iteration: For $l = 0, 1, \ldots, L-1$:

- compute $x^{k,l+1}$ by solving (10) for $y = y^{k,l}$;
- compute $y^{k,l+1}$ from (13) for $x = x^{k,l+1}$;

Output: $x^k = x^{k,L}$, $y^k = y^{k,L}$.

Here L can be chosen using some convergence test techniques. In fact, setting $L = 1$ is sufficient to establish the convergence of the sequence Wu and Tai (2010) generated by Algorithm 0.1. In this case, the augmented Lagrangian method is well known as the alternating direction method of multipliers Boyd (2010).

Convergence analysis

In this section, we present some convergence results of Algorithm 0.1. Actually, we can verify that Algorithm 0.1 is converged in two cases, i.e., when the minimization problem (9) is exactly solved in each iteration and the minimization problem (9) is roughly solved in each iteration Glowinski (1984); Glowinski and Tallec (1989); Wu and Tai (2010). We comment that the convergence proof in Wu and Tai (2010) is based on Glowinski (1984); Glowinski and Tallec (1989), but reduces the uniform convexity assumption of $R(\cdot)$. Here, we just take the main convergence results from Wu and Tai (2010) and omit the details.

In the first case, we should set $L \to \infty$ in Algorithm 0.2 and the inner iteration is guaranteed to converge.

Theorem 3. *The sequence* $\{(x^{k,l}, y^{k,l}) : l = 0, 1, 2, \ldots\}$ *generated by Algorithm* 0.2 *converges to a solution of the problem* (9)*.*

Theorem 4. *Assume that* $(x^*, y^*; \lambda^*)$ *is a saddle-point of* $\mathscr{L}_{\mathrm{TV}}(x, y; \lambda)$. *Suppose that the minimization problem* (9) *is exactly solved in each iteration; i.e.,* $L \to \infty$ *in Algorithm* 0.2. *Then the sequence* $(x^k, y^k; \lambda^k)$ *generated by Algorithm* 0.1 *satisfies*

$$\begin{cases} \lim\limits_{k\to\infty} G_{\mathrm{TV}}(x^k, y^k) = G_{\mathrm{TV}}(x^*, y^*), \\ \lim\limits_{k\to\infty} \|y^k - \nabla x^k\| = 0. \end{cases} \tag{16}$$

Since $R(y)$ *is continuous,* (16) *indicates that* x^k *is a minimizing sequence of* $E_{\mathrm{TV}}(\cdot)$. *If we further have* $\mathrm{Null}(K) = \{0\}$, *then*

$$\begin{cases} \lim\limits_{k\to\infty} x^k = x^*, \\ \lim\limits_{k\to\infty} y^k = y^*. \end{cases}$$

In the second case, we set $L = 1$ in Algorithm 0.2.

Theorem 5. *Assume that* $(x^*, y^*; \lambda^*)$ *is a saddle-point of* $\mathscr{L}_{\mathrm{TV}}(x, y; \lambda)$. *Suppose that the minimization problem* (9) *is roughly solved in each iteration, i.e., with* $L = 1$ *in Algorithm* 0.2. *Then the sequence* $(x^k, y^k; \lambda^k)$ *generated by Algorithm* 0.1 *satisfies*

$$\begin{cases} \lim\limits_{k\to\infty} G_{\mathrm{TV}}(x^k, y^k) = G_{\mathrm{TV}}(x^*, y^*), \\ \lim\limits_{k\to\infty} \|y^k - \nabla x^k\| = 0. \end{cases} \tag{17}$$

Since $R(y)$ *is continuous,* (17) *indicates that* x^k *is a minimizing sequence of* $E_{\mathrm{TV}}(\cdot)$. *If we further have* $\mathrm{Null}(K) = \{0\}$, *then*

$$\begin{cases} \lim\limits_{k\to\infty} x^k = x^*, \\ \lim\limits_{k\to\infty} y^k = y^*. \end{cases}$$

Augmented Lagrangian method for TV-L^2 *restoration with box constraint*

In this section, we review the augmented Lagrangian method for the TV restoration model with the L^2 fidelity term and the box constraint Chan et al (2013), which reads

$$\min_{x\in\mathscr{X}} \left\{ E_{\mathrm{TVB}}(x) = \frac{\alpha}{2}\|Kx - d\|^2 + R(\nabla x) + B(x) \right\}, \tag{18}$$

where $\alpha > 0$, $R(\nabla x)$ is defined as (3), and we have $-\infty < \underline{b} \le \bar{b} < +\infty$ in $B(x)$. This model is also a special case of model (2), where $F(Kx) = \frac{\alpha}{2}\|Kx - d\|^2$.

The box constraint is inherent in digital image processing. The nature image is stored as discrete numerical arrays in some digital media. The typical used ranges are $[0, 1]$ and $[0, 255]$. It has been show that adding the box constraint in image

restoration can improve the quality of the recovered image Beck and Teboulle (2009); Chan and Ma (2012); Chan et al (2013).

The original method proposed in Chan et al (2013) is under the framework of the alternating direction method of multipliers, which is a special case of the augmented Lagrangian method. For the sake of clarity, we reformulate it in our notations and styles.

Compared with the TV-L^2 model (5), this model has one more non-differentiability term $B(x)$. Thus, we need another variable to eliminate the nondifferentiation for x. We introduce two auxiliary variables $y \in \mathscr{Y}$ and $z \in \mathscr{X}$ and rewrite the problem (18) to be the following constrained optimization problem

$$\begin{aligned} \min_{x\in\mathscr{X},y\in\mathscr{Y},z\in\mathscr{X}} & \left\{ G_{\mathrm{TVB}}(x,y,z) = \frac{\alpha}{2}\|Kx-d\|^2 + R(y) + B(z) \right\} \\ \text{s.t.} \quad & \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x, \end{aligned} \tag{19}$$

where $\mathscr{I}_1 : \mathscr{X} \to \mathscr{X}$ is the identify operator.

We define the augmented Lagrangian function for the problem (19) as follows

$$\begin{aligned} \mathscr{L}_{\mathrm{TVB}}(x,y,z;\lambda_y,\lambda_z) = & \frac{\alpha}{2}\|Kx-d\|^2 + R(y) + B(z) \\ & + \left\langle \begin{pmatrix} \lambda_y \\ \lambda_z \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x \right\rangle \\ & + \frac{1}{2}\left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x \right\|^2_{\mathscr{S}}, \end{aligned} \tag{20}$$

where $\begin{pmatrix} \lambda_y \\ \lambda_z \end{pmatrix}$ is the Lagrangian multiplier and $\mathscr{S} = \begin{pmatrix} \beta_y \mathscr{I}_2 & \\ & \beta_z \mathscr{I}_1 \end{pmatrix}$ with the identify operator $\mathscr{I}_2 : \mathscr{Y} \to \mathscr{Y}$ and positive parameters β_y, β_z. Here $\|u\|_{\mathscr{S}}$ denotes the $\mathscr{S}$-norm, defined by $\|u\|_{\mathscr{S}} = \sqrt{\langle u, \mathscr{S}u\rangle}$.

For the augmented Lagrangian method, we consider the saddle-point problem

$$\begin{aligned} & \text{Find } (x^*,y^*,z^*,\lambda_y^*,\lambda_z^*) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{X}\times\mathscr{Y}\times\mathscr{X}, \\ & \text{s.t.} \quad \mathscr{L}_{\mathrm{TVB}}(x^*,y^*,z^*;\lambda_y,\lambda_z) \le \mathscr{L}_{\mathrm{TVB}}(x^*,y^*,z^*;\lambda_y^*,\lambda_z^*) \le \mathscr{L}_{\mathrm{TVB}}(x,y,z;\lambda_y^*,\lambda_z^*), \\ & \qquad \forall (x,y,z,\lambda_y,\lambda_z) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{X}\times\mathscr{Y}\times\mathscr{X}. \end{aligned} \tag{21}$$

Finally, we use an alternating direction iterative scheme in the augmented Lagrangian method to solve the saddle-point problem (21); See Algorithm 0.3.

To solve the minimization problem (22), we separate it into two subproblems respect to x and $\begin{pmatrix} y \\ z \end{pmatrix}$ and employ an alternative minimization procedure.

Algorithm 0.3: Augmented Lagrangian method for TV-L^2 model with box constraint

Initialization: $x^{-1}=0$, $\begin{pmatrix} y^{-1} \\ z^{-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} \lambda_y^0 \\ \lambda_z^0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;

Iteration: For $k=0,1,\ldots$:

1. compute (x^k,y^k,z^k) as an (approximate) minimizer of the augmented Lagrangian functional with the Lagrange multiplier $\begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}$, i.e.,

$$(x^k,y^k,z^k) \approx \arg\min_{(x,y,z)\in\mathscr{X}\times\mathscr{Y}\times\mathscr{Z}} \mathscr{L}_{\mathrm{TVB}}(x,y,z;\lambda_y^k,\lambda_z^k), \tag{22}$$

   where $\mathscr{L}_{\mathrm{TVB}}(x,y,z;\lambda_y^k,\lambda_z^k)$ is as in (20);
2. update

$$\begin{pmatrix} \lambda_y^{k+1} \\ \lambda_z^{k+1} \end{pmatrix} = \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix} + \begin{pmatrix} \beta_y(y^k-\nabla x^k) \\ \beta_z(z^k-x^k) \end{pmatrix}.$$

The solution to sub-minimization problem w.r.t. x

Given $\begin{pmatrix} y \\ z \end{pmatrix}$, the minimization problem (22) with respect to x reads

$$\min_{x\in\mathscr{X}} \left\{ \frac{\alpha}{2}\|Kx-d\|^2 - \left\langle \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}, \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x \right\rangle + \frac{1}{2}\left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x \right\|_{\mathscr{S}}^2 \right\}, \tag{23}$$

whose first-order optimization condition gives a linear equation

$$(\alpha K^*K - \beta_y\Delta + \beta_z\mathscr{I}_1)x = \alpha K^*d - \mathrm{div}(\lambda_y^k+\beta_y y) + \lambda_z^k + \beta_z z. \tag{24}$$

Similar to the equation (10), the above equation can be efficiently solved by fast linear solvers such as FFT and CG.

The solution to sub-minimization problem w.r.t. (y,z)

Given x, the minimization problem (22) with respect to $\begin{pmatrix} y \\ z \end{pmatrix}$ reads

$$\min_{(y,z)\in\mathscr{Y}\times\mathscr{Z}} \left\{ R(y)+B(z) + \left\langle \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + \frac{1}{2}\left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ \mathscr{I}_1 \end{pmatrix} x \right\|_{\mathscr{S}}^2 \right\}, \tag{25}$$

which can be separated into two independent minimization problems:

- y-subproblem:

$$\min_{y\in\mathscr{Y}} \left\{ R(y) + (\lambda_y^k, y) + \frac{\beta_y}{2}\|y-\nabla x\|^2 \right\}, \tag{26}$$

- z-subproblem:

$$\min_{z\in\mathscr{X}}\left\{B(z)+(\lambda_z^k,z)+\frac{\beta_z}{2}\|z-x\|^2\right\}. \tag{27}$$

We can obtain the minimizer of (26) from (13) and the minimizer of (27) as follows

$$z_{i,j}=\mathscr{P}_{[\underline{b},\bar{b}]}(\xi_{i,j}),\ \forall i,j, \tag{28}$$

where $\mathscr{P}_{[\underline{b},\bar{b}]}(\cdot)$ is the projection onto the interval $[\underline{b},\bar{b}]$ and

$$\xi=x-\frac{\lambda_z}{\beta_z}\in\mathscr{X}.$$

After knowing the solutions of the subproblems (23) and (25), we use the following alternative minimization procedure to solve (22); See Algorithm 0.4.

Algorithm 0.4: Augmented Lagrangian method for TV-L^2 model with box constraint – solve the minimization problem (22)

Initialization: $x^{k,0}=x^{k-1}$, $\begin{pmatrix}y^{k,0}\\z^{k,0}\end{pmatrix}=\begin{pmatrix}y^{k-1}\\z^{k-1}\end{pmatrix}$;

Iteration: For $l=0,1,2,...,L-1$:

- compute $x^{k,l+1}$ by solving (24) for $\begin{pmatrix}y\\z\end{pmatrix}=\begin{pmatrix}y^{k,l}\\z^{k,l}\end{pmatrix}$;
- compute $\begin{pmatrix}y^{k,l+1}\\z^{k,l+1}\end{pmatrix}$ from (13) and (28) for $x=x^{k,l+1}$;

Output: $x^k=x^{k,L}$, $\begin{pmatrix}y^k\\z^k\end{pmatrix}=\begin{pmatrix}y^{k,L}\\z^{k,L}\end{pmatrix}$.

The convergence results of Algorithm 0.3 and Algorithm 0.4 are similar to the convergence results proposed in previous section, one can refer to Chan et al (2013) for details.

Augmented Lagrangian method for TV *restoration with non-quadratic fidelity*

In this section, we review the augmented Lagrangian method proposed in Wu et al (2011) for the TV restoration model with non-quadratic fidelity which reads

$$\min_{x\in\mathscr{X}}\{E_{\mathrm{TVNQ}}(x)=R(\nabla x)+F(Kx)\}. \tag{29}$$

where $R(\nabla x)$ is defined as (3). Here, we consider the non-quadratic fidelity $F(Kx)$ which arises for removing non-Gaussian type noises, such as impulsive noise and

Poisson noise. For impulsive noise removal, we usually use the L^1 fidelity Nikolova (2002, 2004)

$$F(Kx) = \alpha \|Kx - d\|_{L^1}, \tag{30}$$

and for Poisson noise removal, we commonly choose the Kullback-Leibler (KL) divergence fidelity Le et al (2007); Brune et al (2009)

$$F(Kx) = \begin{cases} \alpha \sum\limits_{1\le i,j\le N} ((Kx)_{i,j} - d_{i,j}\log(Kx)_{i,j}),\ (Kx)_{i,j} > 0, \forall\, i,j, \\ +\infty, \qquad\qquad \text{otherwise.} \end{cases} \tag{31}$$

In this section, we focus on the augmented Lagrangian method for image restoration with these two non-quadratic fidelities. For other non-quadratic fidelities, one can extend our method accordingly.

The non-quadratic fidelities (30) and (31) are non-smooth. Adopting the idea to cope with total variation term, we require one more auxiliary variable to remove the nonlinearity arising from $F(Kx)$. We first introduce two auxiliary variables y and z and reformulate (29) to an equivalence constrained optimization problem

$$\begin{aligned} &\min_{x\in\mathscr{X}, y\in\mathscr{Y}, z\in\mathscr{X}} \{G_{\mathrm{TVNQ}}(y,z) = R(y) + F(z)\} \\ &\text{s.t.} \quad \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \nabla \\ K \end{pmatrix} x. \end{aligned} \tag{32}$$

We then define the augmented Lagrangian function for (32) as

$$\begin{aligned} \mathscr{L}_{\mathrm{TVNQ}}(x,y,z;\lambda_y,\lambda_z) = R(y) + F(z) \\ + \left\langle \begin{pmatrix} \lambda_y \\ \lambda_z \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ K \end{pmatrix} x \right\rangle + \left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ K \end{pmatrix} x \right\|_{\mathscr{S}}^2 \end{aligned} \tag{33}$$

with Lagrange multiplier $\begin{pmatrix} \lambda_y \\ \lambda_z \end{pmatrix}$ and $\mathscr{S} = \begin{pmatrix} \beta_y \mathscr{I}_2 & \\ & \beta_z \mathscr{I}_1 \end{pmatrix}$, and consider the saddle-point problem

$$\begin{aligned} &\text{Find } (x^*, y^*, z^*, \lambda_y^*, \lambda_z^*) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{X}\times\mathscr{Y}\times\mathscr{X}, \\ &\text{s.t.} \quad \mathscr{L}_{\mathrm{TVNQ}}(x^*,y^*,z^*;\lambda_y,\lambda_z) \le \mathscr{L}_{\mathrm{TVNQ}}(x^*,y^*,z^*;\lambda_y^*,\lambda_z^*) \le \mathscr{L}_{\mathrm{TVNQ}}(x,y,z;\lambda_y^*,\lambda_z^*), \\ &\quad \forall (x,y,z,\lambda_y,\lambda_z) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{X}\times\mathscr{Y}\times\mathscr{X}. \end{aligned} \tag{34}$$

Finally, we use the following iterative algorithm to solve the saddle-point problem (34); See Algorithm 0.5.

We employ an alternating minimization procedure to solve the problem (35).

The solution to sub-minimization problem w.r.t. x

Given $\begin{pmatrix} y \\ z \end{pmatrix}$, we have the subproblem of x as follows

Algorithm 0.5: Augmented Lagrangian method for TV restoration with non-quadratic fidelity

Initialization: $x^{-1}=0$, $\begin{pmatrix} y^{-1} \\ z^{-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} \lambda_y^0 \\ \lambda_z^0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;

Iteration: For $k=0,1,\ldots$:

1. compute (x^k,y^k,z^k) as an (approximate) minimizer of the augmented Lagrangian functional with the Lagrange multipliers $\begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}$, i.e.,

$$(x^k,y^k,z^k) \approx \arg \min_{(x,y,z)\in \mathscr{X}\times\mathscr{Y}\times\mathscr{X}} \mathscr{L}_{\mathrm{TVNQ}}(x,y,z;\lambda_y^k,\lambda_z^k), \tag{35}$$

   where $\mathscr{L}_{\mathrm{TVNQ}}(x,y,z;\lambda_z^k,\lambda_z^k)$ is as in (33);
2. update

$$\begin{pmatrix} \lambda_y^{k+1} \\ \lambda_z^{k+1} \end{pmatrix} = \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix} + \begin{pmatrix} \beta_y(y^k-\nabla x^k) \\ \beta_z(z^k-Kx^k) \end{pmatrix}.$$

$$\min_{x\in\mathscr{X}} \left\{ -\left\langle \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}, \begin{pmatrix} \nabla \\ K \end{pmatrix} x \right\rangle + \left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ K \end{pmatrix} x \right\|_{\mathscr{S}}^2 \right\}, \tag{36}$$

which has the optimization condition

$$(\beta_z K^*K - \beta_y \Delta)x = K^*(\lambda_z^k + \beta_z z) - \mathrm{div}(\lambda_y^k + \beta_y y). \tag{37}$$

We can use fast linear solvers to solve the above equation, such as FFT and CG.

The solution to sub-minimization problem w.r.t. (y,z)

Given x, we have the subproblem of $\begin{pmatrix} y \\ z \end{pmatrix}$ as follows

$$\min_{(y,z)\in\mathscr{Y}\times\mathscr{X}} \left\{ R(y)+F(z)+\left\langle \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + \left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla \\ K \end{pmatrix} x \right\|_{\mathscr{S}}^2 \right\}. \tag{38}$$

We can split it into two distinct minimization problems with respect to y and z as follows

- y-subproblem:

$$\min_{y\in\mathscr{Y}} \left\{ R(y) + (\lambda_y^k, y) + \frac{\beta_y}{2}\|y-\nabla x\|^2 \right\}. \tag{39}$$

- z-subproblem:

$$\min_{z\in\mathscr{X}} \left\{ F(z) + (\lambda_z^k, z) + \frac{\beta_z}{2}\|z-Kx\|^2 \right\}. \tag{40}$$

For the problem (39), it is same as the problem (11) and can be solved via (13). For the problem (40), we next show its solution based on the choices of $F(\cdot)$.

For the L^1 fidelity (30), we can rewrite the z-subproblem (40) as

$$\min_{z\in\mathscr{X}}\left\{\alpha\|z-d\|_{L^1}+\frac{\beta_z}{2}\|z-\xi\|^2\right\}$$

where

$$\xi = Kx-\frac{\lambda_z^k}{\beta_z}.$$

It has closed form solution Donoho (1995); Yang et al (2009b); Wu et al (2011)

$$z_{i,j} = d_{i,j}+\max\left(0,1-\frac{\alpha}{\beta_z|\xi_{i,j}-d_{i,j}|}\right)(\xi_{i,j}-d_{i,j}), \tag{41}$$

which is a 1-dimension case of (13). In this case, the alternating minimization procedure to solve the problem (35) is described in Algorithm 0.6.

Algorithm 0.6: Augmented Lagrangian method for TV restoration with the L^1 fidelity – solve the minimization problem (35)

Initialization: $x^{k,0}=x^{k-1}, \begin{pmatrix} y^{k,0}\\ z^{k,0}\end{pmatrix}=\begin{pmatrix} y^{k-1}\\ z^{k-1}\end{pmatrix}$;

Iteration: For $l=0,1,2,...,L-1$

- compute $x^{k,l+1}$ by solving (37) for $\begin{pmatrix} y\\ z\end{pmatrix}=\begin{pmatrix} y^{k,l}\\ z^{k,l}\end{pmatrix}$;
- compute $\begin{pmatrix} y^{k,l+1}\\ z^{k,l+1}\end{pmatrix}$ from (13) and (41) for $x=x^{k,l+1}$;

Output: $x^k=x^{k,L}, \begin{pmatrix} y^k\\ z^k\end{pmatrix}=\begin{pmatrix} y^{k,L}\\ z^{k,L}\end{pmatrix}$.

For the KL divergence fidelity (31), we can rewrite the z-subproblem (40) as

$$\min_{\substack{z\in\mathscr{X}\\ z_{i,j}>0,\forall i,j}}\left\{\alpha\sum_{1\le i,j\le N}(z_{i,j}-d_{i,j}\log z_{i,j})+\frac{\beta_z}{2}\sum_{1\le i,j\le N}\left|z_{i,j}-\left(Kx-\frac{\lambda_z^k}{\beta_z}\right)_{i,j}\right|^2\right\}.$$

It has closed form solution Setzer et al (2010); Wu et al (2011)

$$z_{i,j}=\frac{1}{2}\left(\sqrt{\left(\xi_{i,j}-\frac{\alpha}{\beta_z}\right)^2+4\frac{\alpha}{\beta_z}d_{i,j}}+\left(\xi_{i,j}-\frac{\alpha}{\beta_z}\right)\right), \tag{42}$$

where

$$\xi = Kx - \frac{\lambda_z^k}{\beta_z}.$$

Now, the alternating minimization procedure to solve the problem (35) with the KL divergence fidelity (31) can be described in Algorithm 0.7.

Algorithm 0.7: Augmented Lagrangian method for TV restoration with the KL divergence fidelity – solve the minimization problem (35)

Initialization: $x^{k,0} = x^{k-1}, \begin{pmatrix} y^{k,0} \\ z^{k,0} \end{pmatrix} = \begin{pmatrix} y^{k-1} \\ z^{k-1} \end{pmatrix}$;

Iteration: For $l = 0, 1, 2, ..., L-1$

- compute $x^{k,l+1}$ from (37) for $\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} y^{k,l} \\ z^{k,l} \end{pmatrix}$;
- compute $\begin{pmatrix} y^{k,l+1} \\ z^{k,l+1} \end{pmatrix}$ from (13) and (42) for $x = x^{k,l+1}$;

Output: $x^k = x^{k,L}, \begin{pmatrix} y^k \\ z^k \end{pmatrix} = \begin{pmatrix} y^{k,L} \\ z^{k,L} \end{pmatrix}$.

The convergence results of Algorithm 0.5, Algorithm 0.6 and Algorithm 0.7 are established in Wu et al (2011), which are similar to convergence results proposed previously for Algorithm 0.1 and Algorithm 0.2.

Extension to multichannel image restoration

In this section, we review the augmented Lagrangian method for the multichannel TV restoration Wu and Tai (2010). The multichannel images are widely used, such as three-channel RGB color image.

The multichannel TV restoration model

We denote an M-channel image by $\boldsymbol{x} = (x_1, x_2, \ldots, x_M)$, where $x_m \in \mathscr{X}$, $\forall m = 1, 2, \ldots, M$. We mention that, at each pixel (i, j), the intensity of $\boldsymbol{x}$ is vector-values, i.e.,

$$\boldsymbol{x}_{i,j} = ((x_1)_{i,j}, (x_2)_{i,j}, \ldots, (x_M)_{i,j}).$$

Let us define

$$\boldsymbol{\mathscr{X}} = \underbrace{\mathscr{X} \times \mathscr{X} \times \cdots \times \mathscr{X}}_{M}, \ \boldsymbol{\mathscr{Y}} = \underbrace{\mathscr{Y} \times \mathscr{Y} \times \cdots \times \mathscr{Y}}_{M}.$$

Then we have $\boldsymbol{x} \in \mathscr{X}$ and

$$\nabla \boldsymbol{x} = (\nabla x_1, \nabla x_2, \ldots, \nabla x_M) \in \mathscr{Y}.$$

The usual inner products and L^2 norms in the spaces $\mathscr{X}$ and $\mathscr{Y}$ are as follows. We denote

$$\langle \boldsymbol{x}, \boldsymbol{z} \rangle = \sum_{1 \le m \le M} \langle x_m, z_m \rangle, \ \ \|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle};$$
$$\langle \boldsymbol{y}, \boldsymbol{w} \rangle = \sum_{1 \le m \le M} \langle y_m, w_m \rangle, \ \|\boldsymbol{y}\| = \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle}.$$

for $\boldsymbol{x}, \boldsymbol{z} \in \mathscr{X}$ and $\boldsymbol{y}, \boldsymbol{w} \in \mathscr{Y}$. At each pixel (i, j), we also define the following pixel-by-pixel norms

$$|\boldsymbol{x}_{i,j}| = \sqrt{\sum_{1 \le m \le M} (x_m)_{i,j}^2} \text{ and } |\boldsymbol{y}_{i,j}| = \sqrt{\sum_{1 \le m \le M} |(y_m)_{i,j}|^2}.$$

for $\boldsymbol{x} \in \mathscr{X}$ and $\boldsymbol{y} \in \mathscr{Y}$.

With reference to the degradation model (1) of the gray image, here we model the multichannel image degradation procedure as

$$\underline{\boldsymbol{x}} \xrightarrow{\text{linear transformation}} \boldsymbol{K}\underline{\boldsymbol{x}} \xrightarrow{\text{noise}} \boldsymbol{d},$$

where $\boldsymbol{d} \in \mathscr{X}$ is an observed image and $\boldsymbol{K} : \mathscr{X} \to \mathscr{X}$ is linear operator like a blur. Here the noise could be also Gaussian, impulsive, Poisson or even others.

In this survey, we consider $\boldsymbol{K}$ as the blur operator and the noise is Gaussian type. The operator $\boldsymbol{K}$ has the form of

$$\boldsymbol{K} = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1M} \\ K_{21} & K_{22} & \cdots & K_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ K_{M1} & K_{M2} & \cdots & K_{MM} \end{pmatrix},$$

where each $K_{i,j}$ is a convolution matrix. The diagonal elements of $\boldsymbol{K}$ denote within-channel blurs, while the off-diagonal elements describe cross-channel blurs. To solve $\underline{\boldsymbol{x}}$, we consider the following multichannel image restoration model Bresson and Chan (2008); Wu and Tai (2010)

$$\min_{\boldsymbol{x} \in \mathscr{X}} \left\{ E_{\text{MTV}}(\boldsymbol{x}) = \frac{\alpha}{2} \|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{d}\|^2 + R_{\text{MTV}}(\nabla \boldsymbol{x}) \right\}, \tag{43}$$

where

$$R_{\text{MTV}}(\nabla \boldsymbol{x}) = \text{TV}(\boldsymbol{x}) = \sum_{1 \le i,j \le N} \sqrt{\sum_{1 \le m \le M} |(\nabla x_m)_{i,j}|^2}$$

is the vectorial TV semi-norm Sapiro and Ringach (1996); Bresson and Chan (2008) (seen Blomgren and Chan (1998) for some other choices).

Similarly as for the single-channel image restoration model, here we make the following assumption:

- $\text{Null}(\nabla) \cap \text{Null}(\boldsymbol{K}) = \{0\}$.

Under this assumption, one can verify that the functional $E_{\text{MTV}}(\boldsymbol{x})$ in (43) is convex, proper, coercive, and lower semi-continuous. Hence, we have the following result Wu and Tai (2010).

Theorem 6. *The problem* (43) *has at least one solution* $\boldsymbol{x}$*, which satisfies*

$$0 \in \alpha \boldsymbol{K}^*(\boldsymbol{K}\boldsymbol{x} - \boldsymbol{d}) - \text{div}\, \partial R_{\text{MTV}}(\nabla \boldsymbol{x}),$$

where $\partial R_{\text{MTV}}(\nabla \boldsymbol{x})$ *is the subdifferential of* R_{MTV} *at* $\nabla \boldsymbol{x}$*. Moreover, if* $\text{Null}(\boldsymbol{K}) = \{0\}$*, the minimizer is unique.*

Augmented Lagrangian method for multichannel TV restoration

By introducing a new variable $\boldsymbol{y} = (y_1, y_2, \ldots, y_M) \in \mathscr{Y}$, we first reformulate the minimization problem (43) to the following equivalent constrained optimization problem:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathscr{X}, \boldsymbol{y}\in\mathscr{Y}} & \left\{ G_{\text{MTV}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\alpha}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{d}\|^2 + R_{\text{MTV}}(\boldsymbol{y}) \right\} \\ \text{s.t.} \quad & \boldsymbol{y} = \nabla \boldsymbol{x}. \end{aligned} \tag{44}$$

We then define the augmented Lagrangian functional as

$$\mathscr{L}_{\text{MTV}}(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{\lambda}) = \frac{\alpha}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{d}\|^2 + R_{\text{MTV}}(\boldsymbol{y}) + \langle \boldsymbol{\lambda}, \boldsymbol{y} - \nabla \boldsymbol{x} \rangle + \frac{\beta}{2}\|\boldsymbol{y} - \nabla \boldsymbol{x}\|^2,$$

with the multiplier $\boldsymbol{\lambda} \in \mathscr{Y}$ and a positive constant β. The augmented Lagrangian method aims at seeking a saddle-point of the following problem:

$$\begin{aligned} &\text{Find} \quad (\boldsymbol{x}^*, \boldsymbol{y}^*; \boldsymbol{\lambda}^*) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{Y} \\ &\text{s.t.} \quad \mathscr{L}_{\text{MTV}}(\boldsymbol{x}^*, \boldsymbol{y}^*; \boldsymbol{\lambda}) \le \mathscr{L}_{\text{MTV}}(\boldsymbol{x}^*, \boldsymbol{y}^*; \boldsymbol{\lambda}^*) \le \mathscr{L}_{\text{MTV}}(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{\lambda}^*) \\ &\quad \forall (\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{\lambda}) \in \mathscr{X} \times \mathscr{Y} \times \mathscr{Y}. \end{aligned} \tag{45}$$

Finally, an iterative procedure to solve the problem (45) is described in Algorithm 0.8.

As for the minimization problem (46), we separate it into two subproblems with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$ and minimize them alternatively.

The solution to sub-minimization problem w.r.t. $\boldsymbol{x}$

For a given $\boldsymbol{y}$, there is the following minimization problem of variable $\boldsymbol{x}$

Algorithm 0.8: Augmented Lagrangian method for the multichannel TV model.

Initialization: $\boldsymbol{x}^{-1}=0$, $\boldsymbol{y}^{-1}=0$, $\boldsymbol{\lambda}^0=0$;
Iteration: For $k=0,1,2,\ldots$:
1. compute $(\boldsymbol{x}^k,\boldsymbol{y}^k)$ from

$$(\boldsymbol{x}^k,\boldsymbol{y}^k)\approx \arg\min_{(\boldsymbol{x},\boldsymbol{y})\in(\mathscr{X},\mathscr{Y})} \mathscr{L}_{\mathrm{MTV}}(\boldsymbol{x},\boldsymbol{y};\boldsymbol{\lambda}^k), \tag{46}$$

2. update

$$\boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k+\beta(\boldsymbol{y}^k-\nabla\boldsymbol{x}^k).$$

$$\min_{\boldsymbol{x}\in\mathscr{X}}\left\{\frac{\alpha}{2}\|\boldsymbol{K}\boldsymbol{x}-\boldsymbol{d}\|^2-\langle\boldsymbol{\lambda}^k,\nabla\boldsymbol{x}\rangle+\frac{\beta}{2}\|\boldsymbol{y}-\nabla\boldsymbol{x}\|^2\right\}. \tag{47}$$

Applying Fourier transforms to the optimality condition of the problem (47), we obtain

$$[\alpha\mathscr{F}(\boldsymbol{K}^*)\mathscr{F}(\boldsymbol{K})-\beta\mathscr{F}(\Delta)]\mathscr{F}(\boldsymbol{x})=\alpha\mathscr{F}(\boldsymbol{K}^*)\mathscr{F}(\boldsymbol{d})-\mathscr{F}(\mathrm{div})\mathscr{F}(\boldsymbol{\lambda}^k+\beta\boldsymbol{y}), \tag{48}$$

from which $\mathscr{F}(\boldsymbol{x})$ can be found and then $\boldsymbol{x}$ via an inverse Fourier transform Yang et al (2009a,b); Wu and Tai (2010); Zhang and Wu (2011). Here applying Fourier transform to a block matrix is regarded as applying Fourier transform to each block.

The solution to sub-minimization problem w.r.t. $\boldsymbol{y}$

For a given $\boldsymbol{x}$, there is the following minimization problem of variable $\boldsymbol{y}$

$$\min_{\boldsymbol{y}\in\mathscr{Y}}\{R_{\mathrm{MTV}}(\boldsymbol{y})+\langle\boldsymbol{\lambda}^k,\boldsymbol{y}\rangle+\frac{\beta}{2}\|\boldsymbol{y}-\nabla\boldsymbol{x}\|^2\}.$$

It has the following closed form solution Yang et al (2009a,b); Wu and Tai (2010); Zhang and Wu (2011)

$$\boldsymbol{y}_{i,j}=\max\left(1-\frac{1}{\beta|\boldsymbol{\eta}_{i,j}|},0\right)\boldsymbol{\eta}_{i,j}, \tag{49}$$

where $\boldsymbol{\eta}=\nabla\boldsymbol{x}-\frac{\boldsymbol{\lambda}^k}{\beta}\in\mathscr{Y}$. Indeed, this solution is a high dimensional version of (13), which can be also derived from the geometric method.

According to (48) and (49), we then have an alternating minimization procedure to (46); See Algorithm 0.9.

We remark that the convergence results of Algorithm 0.3 and Algorithm 0.4 can be directly extended for the Algorithm 0.8 and Algorithm 0.9 Wu and Tai (2010) and we omit the details.

Algorithm 0.9: Augmented Lagrangian method for the multichannel TV model—solve the minimization problem (46).

Initialization: $\boldsymbol{x}^{k,0} = \boldsymbol{x}^{k-1}, \boldsymbol{y}^{k,0} = \boldsymbol{y}^{k-1}$;
Iteration: For $l = 0, 1, 2, \ldots, L-1$:

- compute $\boldsymbol{x}^{k,l+1}$ from (48) for $\boldsymbol{y} = \boldsymbol{y}^{k,l}$;
- compute $\boldsymbol{y}^{k,l+1}$ from (49) for $\boldsymbol{x} = \boldsymbol{x}^{k,l+1}$;

Output: $\boldsymbol{x}^k = \boldsymbol{x}^{k,L}, \boldsymbol{y}^k = \boldsymbol{y}^{k,L}$.

Extension to high order models

In this section, we review augmented Lagrangian methods for some high order models, including the Lysaker-Lundervold-Tai model Lysaker et al (2003), the total generalized variation model Bredies et al (2010); Bredies and Valkonen (2011), the Euler's elastic based model Chan et al (2002); Tai et al (2011) and the mean curvature model Zhu and Chan (2012); Zhu et al (2013a).

Augmented Lagrangian method for Lysaker-Lundervold-Tai model

To overcome the staircase effect, Lysaker, Lundervold and Tai (LLT) suggested regularizing the total variation of the gradient and proposed a model based on second order derivatives Lysaker et al (2003). We begin with some notations to establish this model.

Let

$$\widehat{\mathscr{Y}} = \mathscr{X} \times \mathscr{X} \times \mathscr{X} \times \mathscr{X}.$$

We define the discrete Hessian operator

$$\begin{aligned} H : \mathscr{X} &\to \widehat{\mathscr{Y}} \\ x &\to Hx, \end{aligned}$$

with

$$(Hx)_{i,j} = \begin{pmatrix} (\mathring{D}_{11}^{-+}x)_{i,j} & (\mathring{D}_{12}^{++}x)_{i,j} \\ (\mathring{D}_{21}^{++}x)_{i,j} & (\mathring{D}_{22}^{-+}x)_{i,j} \end{pmatrix},$$

where $\mathring{D}_{11}^{-+}$, $\mathring{D}_{12}^{++}$, $\mathring{D}_{21}^{++}$ and $\mathring{D}_{22}^{-+}$ are second order difference operators and given by

$$\begin{aligned} (\mathring{D}_{11}^{-+}x)_{i,j} &:= (\mathring{D}_1^-(\mathring{D}_1^+x))_{i,j}, \\ (\mathring{D}_{12}^{++}x)_{i,j} &:= (\mathring{D}_1^+(\mathring{D}_2^+x))_{i,j}, \\ (\mathring{D}_{21}^{++}x)_{i,j} &:= (\mathring{D}_2^+(\mathring{D}_1^+x))_{i,j}, \\ (\mathring{D}_{22}^{-+}x)_{i,j} &:= (\mathring{D}_2^-(\mathring{D}_2^+x))_{i,j}. \end{aligned}$$

The usual inner product and L^2 norm in the space $\widehat{\mathscr{Y}}$ are as follows. We denote

$$\langle y,w\rangle = \langle y^1,w^1\rangle + \langle y^2,w^2\rangle + \langle y^3,w^3\rangle + \langle y^4,w^4\rangle \text{ and } \|y\| = \sqrt{\langle y,y\rangle},$$

for $y = \begin{pmatrix} y^1 & y^2 \\ y^3 & y^4 \end{pmatrix} \in \widehat{\mathscr{Y}}$ and $w = \begin{pmatrix} w^1 & w^2 \\ w^3 & w^4 \end{pmatrix} \in \widehat{\mathscr{Y}}$. At each pixel (i,j),

$$|y_{ij}| = \sqrt{(y^1)_{ij}^2 + (y^2)_{ij}^2 + (y^3)_{ij}^2 + (y^4)_{ij}^2}$$

is the usual Euclidean norm in $\mathbb{R}^4$. By using the inner products of $\widehat{\mathscr{Y}}$ and $\mathscr{X}$ and the definitions of the finite difference operators, the adjoint operator of H is as follows

$$H^* : \widehat{\mathscr{Y}} \to \mathscr{X}$$
$$y = \begin{pmatrix} y^1 & y^2 \\ y^3 & y^4 \end{pmatrix} \to H^* y,$$

where

$$(H^*y)_{i,j} = (\mathring{D}_{11}^{+-}y^1)_{i,j} + (\mathring{D}_{21}^{--}y^1)_{i,j} + (\mathring{D}_{12}^{--}y^3)_{i,j} + (\mathring{D}_{22}^{+-}y^4)_{i,j},$$

where $\mathring{D}_{11}^{+-}$, $\mathring{D}_{12}^{--}$, $\mathring{D}_{21}^{--}$, and $\mathring{D}_{22}^{+-}$ are second order difference operators.

By regularizing the norm of the discrete Hessian, the LLT model Lysaker et al (2003); Wu and Tai (2010) reads

$$\min_{x\in\mathscr{X}} \left\{ E_{\mathrm{LLT}}(x) = \frac{\alpha}{2}\|Kx-d\|^2 + R_{\mathrm{HO}}(Hx) \right\}, \tag{50}$$

where $\alpha > 0$, $d \in \mathscr{X}$ is the observed image, $K : \mathscr{X} \to \mathscr{X}$ is the blur operator and

$$R_{\mathrm{HO}}(Hx) = \sum_{1\le i,j\le N} |(Hx)_{i,j}|. \tag{51}$$

Similarly as for the total variation restoration model, we make the following assumption:

- $\mathrm{Null}(H) \cap \mathrm{Null}(K) = \{0\}$.

Under this assumption, the functional $E_{\mathrm{LLT}}(x)$ in (50) is convex, proper, coercive, and lower semi-continuous. Hence, we have the following result Wu and Tai (2010).

Theorem 7. *The problem* (50) *has at least one solution x, which satisfies*

$$0 \in \alpha K^*(Kx-d) + H^*\partial R_{\mathrm{HO}}(Hx),$$

where $\partial R_{\mathrm{HO}}(Hx)$ is the subdifferential of R_{HO} at Hx. Moreover, if $\mathrm{Null}(K) = \{0\}$, *the minimizer is unique.*

In the following we review the augmented Lagrangian method proposed in Wu and Tai (2010) to solve (50). We first introduce a new variable $y \in \widehat{\mathscr{Y}}$ and reformulate (50) into a constrained optimization problem

$$\begin{aligned} \min_{x\in\mathscr{X},y\in\widehat{\mathscr{Y}}} & \left\{ G_{\mathrm{LLT}}(x,y) = \frac{\alpha}{2}\|Kx-d\|^2 + R_{\mathrm{HO}}(y) \right\} \\ \text{s.t.} \quad & y = Hx. \end{aligned} \tag{52}$$

To solve (52), we define the augmented Lagrangian functional as

$$\mathscr{L}_{\mathrm{LLT}}(x,y;\lambda) = \frac{\alpha}{2}\|Kx-d\|^2 + R_{\mathrm{HO}}(y) + \langle \lambda, y-Hx\rangle + \frac{\beta}{2}\|y-Hx\|^2, \tag{53}$$

with the multiplier $\lambda \in \widehat{\mathscr{Y}}$ and a positive constant β, and consider the following saddle-point problem:

$$\begin{aligned} &\text{Find } (x^*,y^*,\lambda^*) \in \mathscr{X}\times\widehat{\mathscr{Y}}\times\widehat{\mathscr{Y}} \\ &\text{s.t. } \mathscr{L}_{\mathrm{LLT}}(x^*,y^*;\lambda) \le \mathscr{L}_{\mathrm{LLT}}(x^*,y^*;\lambda^*) \le \mathscr{L}_{\mathrm{LLT}}(x,y;\lambda^*) \\ &\quad \forall (x,y;\lambda) \in \mathscr{X}\times\widehat{\mathscr{Y}}\times\widehat{\mathscr{Y}}. \end{aligned} \tag{54}$$

We employ an iterative procedure to solve the saddle-point problem (54), which is described as Algorithm 0.10.

Algorithm 0.10: Augmented Lagrangian method for the LLT model.

Initialization: $x^{-1}=0$, $y^{-1}=0$,$\lambda^0=0$;
Iteration: For $k=0,1,2,\ldots$:
1. compute (x^k,x^k) from

$$(x^k,x^k) \approx \arg\min_{(x,y)\in(\mathscr{X},\widehat{\mathscr{Y}})} \mathscr{L}_{\mathrm{LLT}}(x,y;\lambda^k), \tag{55}$$

2. update

$$\lambda^{k+1} = \lambda^k + \beta(y^k - Hx^k).$$

The solution to sub-minimization problem w.r.t. x

Given y, we are going to solve the following minimization problem

$$\min_{x\in\mathscr{X}} \left\{ \frac{\alpha}{2}\|Kx-d\|^2 - \langle \lambda^k, Hx\rangle + \frac{\beta}{2}\|y-Hx\|^2 \right\}, \tag{56}$$

the first-order optimal condition of which gives us a linear equation as follows

$$(\alpha K^*K + \beta H^*H)x = \alpha K^*d + H^*(\lambda^k + \beta y). \tag{57}$$

This equation can solved by well-developed linear solvers such as FFT and CG.

The solution to sub-minimization problem w.r.t. y

Given x, we are going to solve the following minimization problem

$$\min_{y\in\widehat{\mathscr{Y}}}\left\{R_{\mathrm{HO}}(y)+(\lambda^k,y)+\frac{\beta}{2}\|y-Hx\|^2\right\}, \tag{58}$$

the closed form solution of which is

$$y_{i,j}=\max\left(0,1-\frac{1}{\beta|\eta_{i,j}|}\right)\eta_{ij}, \tag{59}$$

where $\eta=Hx-\frac{\lambda^k}{\beta}\in\widehat{\mathscr{Y}}$. We mention that the solution (59) is a high dimensional version of (13), which can be also derived from the geometric method.

According to (57) and (59), we then use an iterative procedure to alternatively calculate x and y ; see Algorithm 0.11.

Algorithm 0.11: Augmented Lagrangian method for the LLT model—solve the minimization problem (55).

Initialization: $x^{k,0}=x^{k-1},y^{k,0}=y^{k-1}$;
Iteration: For $l=0,1,2,\ldots,L-1$:

- compute $x^{k,l+1}$ by solving (57) for $y=y^{k,l}$;
- compute $y^{k,l+1}$ from (59) for $x=x^{k,l+1}$;

Output: $x^k=x^{k,L},y^k=y^{k,L}$.

We mention that the convergence results of the augmented Lagrangian method for the LLT model are straightforward as in Wu and Tai (2010) and we omit the details.

Augmented Lagrangian method for total generalized variation model

Total generalized variation (TGV) is a very successful generalization of total variation, which involves high order derivatives to reduce staircase effect Bredies et al (2010); Bredies and Valkonen (2011). In this section, we consider the following discrete second-order total generalized variation Bredies et al (2010); Bredies and Valkonen (2011) based image restoration model

$$\min_{x\in\mathscr{X},w\in\mathscr{Y}}\left\{\frac{1}{2}\|Kx-d\|^2+\alpha_1 R(\nabla x-w)+\alpha_0 R_{\mathrm{HO}}(\mathscr{E}w)\right\}, \tag{60}$$

where $R(\nabla x-w)$ is defined by replacing ∇x by $\nabla x-w$ in (3), $\mathscr{E}$ denotes a distributional symmetrized gradient operator

$$\begin{aligned}\mathscr{E}&:\mathscr{Y}\to\widehat{\mathscr{Y}}\\ w=(w^1,w^2)&\to \mathscr{E}w=\frac{1}{2}(\nabla w+(\nabla w)^T),\end{aligned}$$

with

$$\begin{aligned}(\mathscr{E}w)_{ij}&=\frac{1}{2}(\nabla w+(\nabla w)^T)_{ij}\\ &=\begin{pmatrix}(\mathring{D}_1^+w^1)_{ij} & \frac{1}{2}((\mathring{D}_2^+w^1)_{ij}+(\mathring{D}_1^+w^2)_{ij})\\ \frac{1}{2}((\mathring{D}_2^+w^1)_{ij}+(\mathring{D}_1^+w^2)_{ij}) & (\mathring{D}_2^+w^2)_{ij}\end{pmatrix},\end{aligned}$$

and $R_{\mathrm{HO}}(\cdot)$ is defined in (51). Similarly, by using the inner products of $\widehat{\mathscr{Y}}$ and $\mathscr{Y}$ and the definitions of the finite difference operators the adjoint operator of $-\mathscr{E}$ is as follows

$$\begin{aligned}\mathrm{div}_2&:\widehat{\mathscr{Y}}\to\mathscr{Y}\\ z=\begin{pmatrix}z^1 & z^3\\ z^3 & z^2\end{pmatrix}&\to \mathrm{div}_2 z,\end{aligned}$$

where

$$\mathrm{div}_2 z=\begin{pmatrix}\mathring{D}_1^-z^1+\mathring{D}_2^-z^3\\ \mathring{D}_1^-z^3+\mathring{D}_2^-z^2\end{pmatrix}$$

with

$$(\mathrm{div}_2 z)_{ij}=\begin{pmatrix}(\mathring{D}_1^-z^1)_{ij}+(\mathring{D}_2^-z^3)_{ij}\\ (\mathring{D}_1^-z^3)_{ij}+(\mathring{D}_2^-z^2)_{ij}\end{pmatrix}.$$

Augmented Lagrangian based methods for total generalized variation related models can be found in Shirai and Okuda (2014); Liu (2016); Gao et al (2018). Here, we propose the augmented Lagrangian method to solve (60). We first introduce two auxiliary variable $y=(y^1,y^2)\in\mathscr{Y}$ and $z=\begin{pmatrix}z^1 & z^3\\ z^3 & z^2\end{pmatrix}\in\widehat{\mathscr{Y}}$ and transform it into an equivalent constrained optimization problem

$$\begin{aligned}\min_{x\in\mathscr{X},w\in\mathscr{Y},y\in\mathscr{Y},z\in\widehat{\mathscr{Y}}}&\left\{G_{\mathrm{TGV}}(x,y,z)=\frac{1}{2}\|Kx-d\|^2+\alpha_1 R(y)+\alpha_0 R_{\mathrm{HO}}(z)\right\}\\ \text{s.t.}\quad&\begin{pmatrix}y\\ z\end{pmatrix}=\begin{pmatrix}\nabla & -\mathscr{I}_2\\ & \mathscr{E}\end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}.\end{aligned} \tag{61}$$

We then define the augmented Lagrangian function as follows

$$\begin{aligned}\mathscr{L}_{\mathrm{TGV}}(x,w,y,z;\lambda_y,\lambda_z) = \frac{1}{2}\|Kx-d\|^2 + \alpha_1 R(y) + \alpha_0 R_{\mathrm{HO}}(z) \\ + \left\langle \begin{pmatrix}\lambda_y\\ \lambda_z\end{pmatrix}, \begin{pmatrix}y\\ z\end{pmatrix} - \begin{pmatrix}\nabla & -\mathscr{I}_2\\ & \mathscr{E}\end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}\right\rangle \\ + \frac{1}{2}\left\| \begin{pmatrix}y\\ z\end{pmatrix} - \begin{pmatrix}\nabla & -\mathscr{I}_2\\ & \mathscr{E}\end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}\right\|^2_{\mathscr{S}},\end{aligned} \tag{62}$$

where $\begin{pmatrix}\lambda_y\\ \lambda_z\end{pmatrix}$ is the Lagrange multiplier and $\mathscr{S} = \begin{pmatrix}\beta_y\mathscr{I}_2 & \\ & \beta_z\widehat{\mathscr{I}_2}\end{pmatrix}$ with the identify operator $\widehat{\mathscr{I}_2} : \widehat{\mathscr{Y}} \to \widehat{\mathscr{Y}}$, and consider the saddle point problem

$$\begin{aligned}&\text{Find } (x^*,w^*,y^*,z^*,\lambda_y^*,\lambda_z^*) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{Y}\times\widehat{\mathscr{Y}}\times\mathscr{Y}\times\widehat{\mathscr{Y}}\\ &\text{s.t. } \mathscr{L}_{\mathrm{TGV}}(x^*,w^*,y^*,z^*;\lambda_y,\lambda_z) \le \mathscr{L}_{\mathrm{TGV}}(x^*,w^*,y^*,z^*;\lambda_y^*,\lambda_z^*) \le \mathscr{L}_{\mathrm{TGV}}(x,w,y,z;\lambda_y^*,\lambda_z^*),\\ &\quad \forall (x,w,y,z,\lambda_y,\lambda_z) \in \mathscr{X}\times\mathscr{Y}\times\mathscr{Y}\times\widehat{\mathscr{Y}}\times\mathscr{Y}\times\widehat{\mathscr{Y}}.\end{aligned} \tag{63}$$

Finally, the iterative algorithm for seeking a saddle point is given by Algorithm 0.12.

Algorithm 0.12: Augmented Lagrangian method for TGV model

Initialization: $\begin{pmatrix}x^{-1}\\ w^{-1}\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$, $\begin{pmatrix}y^{-1}\\ z^{-1}\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$, $\begin{pmatrix}\lambda_y^0\\ \lambda_z^0\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$;

Iteration: For $k = 0,1,\ldots$:

1. compute (x^k,w^k,y^k,z^k) from $\begin{pmatrix}\lambda_y^k\\ \lambda_z^k\end{pmatrix}$, i.e.,

$$(x^k,w^k,y^k,z^k) \approx \arg\min_{(x,w,y,z)\in\mathscr{X}\times\mathscr{Y}\times\mathscr{Y}\times\widehat{\mathscr{Y}}} \mathscr{L}_{\mathrm{TGV}}(x,w,y,z;\lambda_y^k,\lambda_z^k), \tag{64}$$

2. update

$$\begin{pmatrix}\lambda_y^{k+1}\\ \lambda_z^{k+1}\end{pmatrix} = \begin{pmatrix}\lambda_y^k\\ \lambda_z^k\end{pmatrix} + \begin{pmatrix}\beta_y(y^k - \nabla x^k + w^k)\\ \beta_z(z^k - \mathscr{E}w^k)\end{pmatrix}.$$

The solution to sub-minimization problem w.r.t. (x,w)

Given $\begin{pmatrix}y\\ z\end{pmatrix}$, we concern with the following minimization problem

$$\min_{(x,w)\in\mathscr{X}\times\mathscr{Y}} \left\{ \frac{1}{2}\|Kx-d\|^2 - \left\langle \begin{pmatrix} \lambda_y^k \\ \lambda_z^k \end{pmatrix}, \begin{pmatrix} \nabla & -\mathscr{I}_2 \\ & \mathscr{E} \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \right\rangle + \frac{1}{2} \left\| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \nabla & -\mathscr{I}_2 \\ & \mathscr{E} \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \right\|_{\mathscr{S}}^2 \right\}. \quad (65)$$

This problem is a quadratic optimization problem, whose optimal condition gives a linear system equations

$$\begin{pmatrix} K^*K-\beta_y\Delta & \beta_y \operatorname{div} \\ -\beta_y\nabla & \beta_y-\beta_z \operatorname{div}_2 \mathscr{E} \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} K^*d-\operatorname{div}(\lambda_y^k+\beta_y y) \\ -\lambda_y^k-\beta_y y-\operatorname{div}_2(\lambda_z^k+\beta_z z) \end{pmatrix},$$

i.e.,

$$\begin{cases} (K^*K-\beta_y\mathring{D}_1^-\mathring{D}_1^+-\beta_y\mathring{D}_2^-\mathring{D}_2^+)x+\beta_y\mathring{D}_1^-w^1+\beta_y\mathring{D}_2^-w^2=g^1, \\ -\beta_y\mathring{D}_1^+x+(\beta_y\mathscr{I}-\beta_z\mathring{D}_1^-\mathring{D}_1^+-\dfrac{\beta_z}{2}\mathring{D}_2^-\mathring{D}_2^+)w^1-\dfrac{\beta_z}{2}\mathring{D}_2^-\mathring{D}_1^+w^2=g^2, \\ -\beta_y\mathring{D}_2^+x-\dfrac{\beta_z}{2}\mathring{D}_1^-\mathring{D}_2^+w^1+(\beta_y\mathscr{I}-\dfrac{\beta_z}{2}\mathring{D}_1^-\mathring{D}_1^+-\beta_z\mathring{D}_2^-\mathring{D}_2^+)w^2=g^3, \end{cases} \quad (66)$$

where

$$\begin{aligned} g^1 &= K^*d-\mathring{D}_1^-\big((\lambda_y^k)^1+\beta_y y^1\big)-\mathring{D}_2^-\big((\lambda_y^k)^2+\beta_y y^2\big), \\ g^2 &= -(\lambda_y^k)^1-\beta_y y^1-\mathring{D}_1^-\big((\lambda_z^k)^1+\beta_z z^1\big)-\mathring{D}_2^-\big((\lambda_z^k)^3+\beta_z z^3\big), \\ g^3 &= -(\lambda_y^k)^2-\beta_y y^2-\mathring{D}_1^-\big((\lambda_z^k)^3+\beta_z z^3\big)-\mathring{D}_2^-\big((\lambda_z^k)^2+\beta_z z^2\big). \end{aligned}$$

This linear system equations with periodic boundary condition can be efficiently solved by Fourier transform via FFT implementation Yang et al (2009a); Zhang and Wu (2011). Firstly, we apply FFTs to both sides of (66),

$$\begin{pmatrix} a^{11} & a^{12} & a^{13} \\ a^{21} & a^{22} & a^{23} \\ a^{31} & a^{32} & a^{33} \end{pmatrix} \begin{pmatrix} \mathscr{F}(x) \\ \mathscr{F}(w^1) \\ \mathscr{F}(w^2) \end{pmatrix} = \begin{pmatrix} \mathscr{F}(g^1) \\ \mathscr{F}(g^2) \\ \mathscr{F}(g^3) \end{pmatrix}. \quad (67)$$

where $a^{i,j},(i,j=1,\dots 3)$ are Fourier coefficients of the operators in the left side of (66). Secondly, we solve the resulting systems by block Gaussian elimination method for $\mathscr{F}(x)$, $\mathscr{F}(w^1)$ and $\mathscr{F}(w^2)$. Finally, we apply inverse FFTs to obtain x and $w=(w^1,w^2)$.

The solution to sub-minimization problem w.r.t. (y,z)

Given $\begin{pmatrix} x \\ w \end{pmatrix}$, we concern with the following minimization problem

$$\min_{(y,z)\in\mathscr{Y}\times\widehat{\mathscr{Y}}}\left\{\alpha_1 R(y)+\alpha_0 R_{\mathrm{HO}}(z)+\left\langle\begin{pmatrix}\lambda_y^k\\ \lambda_z^k\end{pmatrix},\begin{pmatrix}y\\ z\end{pmatrix}\right\rangle + \frac{1}{2}\left\|\begin{pmatrix}y\\ z\end{pmatrix}-\begin{pmatrix}\nabla & -\mathscr{I}_2\\ \mathscr{E} & \end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}\right\|_{\mathscr{S}}^2\right\}. \quad (68)$$

It can be separated into two independent minimization problems:

- y-subproblem:

$$\min_{y\in\mathscr{Y}}\left\{\alpha_1 R(y)+\langle\lambda_y^k,y\rangle+\frac{\beta_y}{2}\|y-\nabla x+w\|^2\right\}, \quad (69)$$

- z-subproblem:

$$\min_{z\in\widehat{\mathscr{Y}}}\left\{\alpha_0 R_{\mathrm{HO}}(z)+\langle\lambda_z^k,z\rangle+\frac{\beta_z}{2}\|z-\mathscr{E}w\|^2\right\}. \quad (70)$$

The problem (69) and (70) have the closed form solutions

$$y_{i,j}=\max\left(0,1-\frac{\alpha_1}{\beta_y|\eta_{i,j}|}\right)\eta_{i,j},\ \text{and}\ z_{i,j}=\max\left(0,1-\frac{\alpha_0}{\beta_z|\xi_{i,j}|}\right)\xi_{i,j}, \quad (71)$$

where

$$\eta=\nabla x-w-\frac{\lambda_y^k}{\beta_y}\in\mathscr{Y},\ \text{and}\ \xi=\mathscr{E}w-\frac{\lambda_z^k}{\beta_z}\in\widehat{\mathscr{Y}}.$$

After knowing the solutions of the subproblems (65) and (68), we use the following alternative minimization procedure to solve (64); See Algorithm 0.13.

Algorithm 0.13: Augmented Lagrangian method for TGV model–solve the minimization problem (64)

Initialization: $\begin{pmatrix}x^{k,0}\\ w^{k,0}\end{pmatrix}=\begin{pmatrix}x^{k-1}\\ w^{k-1}\end{pmatrix}$, $\begin{pmatrix}y^{k,0}\\ z^{k,0}\end{pmatrix}=\begin{pmatrix}y^{k-1}\\ z^{k-1}\end{pmatrix}$;

Iteration: For $l=0,1,\ldots,L-1$:

- compute $\begin{pmatrix}x^{k,l+1}\\ w^{k,l+1}\end{pmatrix}$ from (67) for $\begin{pmatrix}y\\ z\end{pmatrix}=\begin{pmatrix}y^{k,l}\\ z^{k,l}\end{pmatrix}$;
- compute $\begin{pmatrix}y^{k,l+1}\\ z^{k,l+1}\end{pmatrix}$ from (71) for $\begin{pmatrix}x\\ w\end{pmatrix}=\begin{pmatrix}x^{k,l+1}\\ w^{k,l+1}\end{pmatrix}$;

Output: $\begin{pmatrix}x^{k}\\ w^{k}\end{pmatrix}=\begin{pmatrix}x^{k,L}\\ w^{k,L}\end{pmatrix}$, $\begin{pmatrix}y^{k}\\ z^{k}\end{pmatrix}=\begin{pmatrix}y^{k,L}\\ z^{k,L}\end{pmatrix}$.

Augmented Lagrangian method for Euler elastic based model

As basic geometric measurements of curves, both length and curvatures are natural regularities that are widely used in various image processing problems. Euler's elastica is defined as the line energy for a smooth planar curves γ

$$E(\gamma) = \int_{\gamma} (a + b\kappa^2) ds, \tag{72}$$

where κ is the curvature of the curve, s is arc length, a, b are positive constants. By summing up the Euler's elastica energies of all the level sets for an image x, it gives the following energy for image denoising task

$$\min_{x} \; R_{\mathrm{EE}}(\kappa(x), \nabla x) + \frac{1}{2}\|Kx - d\|^2, \tag{73}$$

where $\kappa(x) = \mathrm{div}(\frac{\nabla x}{|\nabla x|})$ and $R_{\mathrm{EE}}(\kappa(x), \nabla x)$ is defined by

$$R_{\mathrm{EE}}(\kappa(x), \nabla x) = \sum_{1 \le i,j \le N} \Big(a + b\kappa^2(x_{i,j})\Big) |(\nabla x)_{i,j}|.$$

Euler's elastica regularization has lots applications in shape and image processing. However, the non-convexity, the non-smoothness and the nonlinearity of the Euler's elastica energy make its minimization a challenging task. Chan, Kang and Shen Chan et al (2002) developed a computational scheme based on numerical PDEs for inpainting problem. Bae, Shi and Tai Bae et al (2010) presented an efficient minimization algorithm based on graph cuts for minimizing the Euler's elastica energy. Tai, Hahn and Chung Tai et al (2011) proposed an augmented Lagrangian method based on the operator splitting and relaxation techniques, which greatly improved the efficiency of the Euler's elastica model. Since then, operator splitting and augmented Lagrangian method have been extensively studied for Euler's elastica Duan et al (2013); Zhang et al (2017b); Yashtini and Kang (2016). Recent advances include functional lifting to get a convex, lower semi-continuous, coercive approximation of the Euler's elastica energy Bredies et al (2015), a lifted convex representation of curvature depending energies in the roto-translational space followed by primal-dual scheme Chambolle and Pock (2019), and a Lie operator-splitting based time discretization scheme Deng et al (2019). In Tai et al (2011), Euler's elastica regularized model (73) is reformulated as the following constrained minimization problem

$$\begin{aligned} \min_{x,y,n,m} \; & R_{\mathrm{EE}}(\mathrm{div}\, n, y) + \frac{1}{2}\|Kx - d\|^2 + I_{\mathscr{M}}(m) \\ \text{s.t.} \quad & y = \nabla x, \; n = m, \; |y| = m \cdot y, \end{aligned} \tag{74}$$

where $I_{\mathscr{M}}(\cdot)$ is an indicator function of the set

$$\mathscr{M} = \{m_{ij} : |m_{i,j}| \le 1,\ \forall\, 1 \le i,j \le N\}.$$

Note that the variable m was introduced to relax the constraint on variable n. By requiring m to be lain in the set $\mathscr{M}$, the term $|y| - y \cdot m$ is guaranteed non-negative, which make the sub-minimization problem w.r.t. m easy to handle with. We can further define the augmented Lagrangian functional as follows

$$\begin{aligned}\mathscr{L}_{\mathrm{EE}}(x,y,n,m;\lambda_y,\lambda_n,\lambda_m) &= R_{\mathrm{EE}}(\operatorname{div} n, y) + \frac{1}{2}\|Kx-d\|^2 + I_{\mathscr{M}}(m)\\ &+ \langle \lambda_y, y - \nabla x\rangle + \frac{\beta_y}{2}\|y-\nabla x\|^2 + \langle \lambda_n, n-m\rangle + \frac{\beta_n}{2}\|n-m\|^2\\ &+ \langle \lambda_m, |y| - m\cdot y\rangle + \langle |y| - m\cdot y, \beta_m\rangle,\end{aligned} \tag{75}$$

where λ_y, λ_n, λ_m are the Lagrange multipliers and β_y, β_n, β_m are positive parameters. The iterative algorithm is used to find a point satisfying the first-order condition; see Algorithm 0.14.

Algorithm 0.14: Augmented Lagrangian method for Euler elastic model

Initialization: $x^{-1} = 0$, $y^{-1} = 0$, $n^{-1} = 0$, $m^{-1} = 0$, $\lambda_y^0 = 0$, $\lambda_n^0 = 0$, $\lambda_m^0 = 0$;
Iteration: For $k = 0, 1, \ldots$:

1. Compute (x^k, y^k, n^k, m^k) from

$$(x^k, y^k, n^k, m^k) \approx \arg\min_{(x,y,n,m)} \mathscr{L}_{\mathrm{EE}}(x,y,n,m;\lambda_y^k,\lambda_n^k,\lambda_m^k), \tag{76}$$

2. Update

$$\begin{pmatrix}\lambda_y^{k+1}\\ \lambda_n^{k+1}\\ \lambda_m^{k+1}\end{pmatrix} = \begin{pmatrix}\lambda_y^{k}\\ \lambda_n^{k}\\ \lambda_m^{k}\end{pmatrix} + \begin{pmatrix}\beta_y(y^k - \nabla x^k)\\ \beta_n(n^k - m^k)\\ \beta_m(|y^k| - m^k\cdot y^k)\end{pmatrix}$$

Before we discuss the solution to the minimization problem (76), we define a staggered grid system in Figure 2; see more details of the implementation in Tai et al (2011). We separate the minimization problem (76) into subproblems to pursue the solutions in an alternative mechanism.

The solution to sub-minimization problem w.r.t. x

Given y, we solve the following minimization problem

$$\min_x \frac{1}{2}\|Kx-d\|^2 + \frac{\beta_y}{2}\|y-\nabla x\|^2 - \langle \lambda_y, \nabla x\rangle, \tag{77}$$

the first-order optimal condition of which gives us

$$(K^*K - \beta_y\Delta)x = K^*d - \beta_y \operatorname{div} y - \operatorname{div}\lambda_y.$$

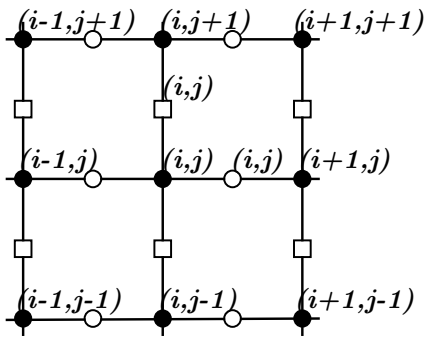


Fig. 2: The rule of indexing variables in the augmented Lagrangian functional (75): x, z, λ_z, λ_m are defined on $\bullet$-nodes. The first and second component of y, n, m, λ_y, λ_n are defined on $\circ$-nodes and $\square$-node, respectively.

Fast numerical methods can be used to solve the above equation such as fast Fourier transform (FFT) and iterative schemes.

The solution to sub-minimization problem w.r.t. y

Given x, n and m, we have the subproblem of y as follows

$$\min_y \ \langle a+b(\operatorname{div} n)^2, |y|\rangle + \langle \lambda_y, y\rangle + \langle \lambda_m+\beta_m, |y| - m\cdot y\rangle + \frac{\beta_y}{2}\|y-\nabla x\|^2, \qquad (78)$$

which can be simplified as

$$\min_y \ \frac{\beta_y}{2}\left\|y-\left(\nabla x+(\frac{\lambda_m+\beta_m}{\beta_y})m-\frac{\lambda_y}{\beta_y}\right)\right\|^2+\left\langle |y|, a+b(\operatorname{div} n)^2+\lambda_m+\beta_m\right\rangle.$$

Such the L^1 regularized minimization problem can be efficiently solved by the closed-form solution.

The solution to sub-minimization problem w.r.t. m

Given n and y, the sub-minimization problem of variable m becomes

$$\min_m \ I_{\mathscr{M}}(m) - \langle \lambda_m, m\rangle + \frac{\beta_n}{2}\|n-m\|^2 - \langle(\lambda_m+\beta_m)y, m\rangle. \qquad (79)$$

We can reformulate the above minimization into a quadratic problem as follows

$$\min_m I_{\mathscr{M}}(m) + \frac{\beta_n}{2}\left\| m - \frac{(\lambda_m + \beta_m)y + \lambda_m}{\beta_n} - n \right\|^2,$$

the optimal solution of which can be achieved by performing the one-step projection to the solution of the quadratic minimization.

The solution to sub-minimization problem w.r.t. n

Given m and y, we are going to solve the following minimization problem of n

$$\min_n \ \langle b(\operatorname{div} n)^2, |y| \rangle + \langle \lambda_n, n \rangle + \frac{\beta_n}{2}\|n-m\|^2, \tag{80}$$

the Euler-Lagrange equation of which is

$$-2\nabla(b|y|\operatorname{div} n) + \beta_n(n-m) + \lambda_n = 0,$$

and can be solved by a frozen coefficient method for easier implementation Tai et al (2011); Yashtini and Kang (2016).

Augmented Lagrangian method for mean curvature-based model

Mean curvature-based model Zhu and Chan (2012) considers an image restoration problem as a surface smoothing task. A basic model is as follows

$$\min_x \ \int_\Omega \left| \operatorname{div}\left(\frac{\nabla x}{\sqrt{1+|\nabla x|^2}} \right) \right| dx + \frac{\alpha}{2}\int_\Omega (Kx-d)^2 dx. \tag{81}$$

Originally, the smoothed mean curvature model (81) was numerically solved by the gradient descent method, which involves high order derivatives and converges slowly in practice. Zhu, Tai and Chan Zhu et al (2013a) developed an augmented Lagrangian method for a mean curvature based image denoising model (81), with similar ideas further studied in Sun and Chen (2014); Myllykoski et al (2015); Zhang (2018). Following Zhu et al (2013a), we rewrite the mean curvature-regularized model into the following constrained minimization problem

$$\begin{aligned} &\min_{x,y,q,n,m} R_{\mathrm{MC}}(q) + \frac{\alpha}{2}\|Kx-d\|^2 + I_{\mathscr{M}}(m) \\ &\text{s.t. } \ y = \langle \nabla x, 1 \rangle, \ q = \operatorname{div} n, \ n = m, \ |y| = y \cdot m, \end{aligned} \tag{82}$$

where $R_{\mathrm{MC}}(q)$ is defined as

$$R_{\mathrm{MC}}(q) = \sum_{1 \le i,j \le N} |q_{i,j}|.$$

The corresponding augmented Lagrangian functional for the constrained minimization problem is defined as

$$\begin{aligned}\mathscr{L}_{\mathrm{MC}}(x,y,q,m,n;\lambda_y,\lambda_q,\lambda_n,\lambda_m) &= R_{\mathrm{MC}}(q)+\frac{\alpha}{2}\|Kx-d\|^2+I_{\mathscr{M}}(m)\\ &+\big\langle\lambda_y,\langle\nabla x,1\rangle\big\rangle+\frac{\beta_y}{2}\|y-\langle\nabla x,1\rangle\|^2+\langle q-\nabla\cdot n\rangle+\frac{\beta_q}{2}\|q-\nabla\cdot n\rangle\\ &+\langle\lambda_n,n-m\rangle+\frac{\beta_n}{2}\|n-m\|^2+\langle\lambda_m,|y|-y\cdot m\rangle+\beta_m\langle|y|-y\cdot m\rangle,\end{aligned}\tag{83}$$

where $\lambda_y,\lambda_q,\lambda_n,\lambda_m$ are Lagrange multipliers, and $\beta_y,\beta_q,\beta_n,\beta_m$ are positive parameters. The iterative algorithm is used to find a point satisfying the first-order condition; see Algorithm 0.15.

Algorithm 0.15: Augmented Lagrangian method for mean curvature-based model

Initialization: $x^{-1}=0$, $y^{-1}=0$, $q^{-1}=0$, $n^{-1}=0$, $m^{-1}=\mathbf{0}$, $\lambda_y^0=0$, $\lambda_q^0=0$, $\lambda_n^0=0$, $\lambda_m^0=0$;

Iteration: For $k=0,1,\ldots$:

1. Compute (x^k,y^k,q^k,n^k,m^k) from

$$(x^k,y^k,q^k,n^k,m^k)\approx\arg\min_{(x,y,q,n,m)}\mathscr{L}_{\mathrm{MC}}(x,y,q,n,m;\lambda_y^k,\lambda_q^k,\lambda_n^k,\lambda_m^k),\tag{84}$$

2. Update

$$\begin{pmatrix}\lambda_y^{k+1}\\ \lambda_q^{k+1}\\ \lambda_n^{k+1}\\ \lambda_m^{k+1}\end{pmatrix}=\begin{pmatrix}\lambda_y^{k}\\ \lambda_q^{k}\\ \lambda_n^{k}\\ \lambda_m^{k}\end{pmatrix}+\begin{pmatrix}\beta_y(y^k-\langle\nabla x^k,1\rangle)\\ \beta_q(q^k-\nabla\cdot n^k)\\ \beta_n(n^k-m^k)\\ \beta_m(|y^k|-y^k\cdot m^k)\end{pmatrix}$$

We can separate the minimization problem (84) into subproblems to obtain the solutions in an alternative way. Similarly as discussed for Euler's elastica model, the minimizers to the variable y, q and m have closed-form solutions, while the minimizers to the variable x and n are obtained by solving the associated Euler-Lagrange equations by either FFT or fast iterative schemes. Therefore, we omit the details here.

Numerical Experiments

In this section, we give some numerical results of augmented Lagrangian methods for solving the total variation related image restoration models. For each model, we test only one image by considering the limit space. For more examples, please refer to literatures Tai and Wu (2009); Wu and Tai (2010); Wu et al (2011); Chan et al (2013); Tai et al (2011); Zhu et al (2013a). We perform the numerical experiments

in MATLAB R2018A (Version 9.4) on a MacBook Pro with 2.3 GHz dual-core Intel Core i5 processor and 8GB memory. For each experiment, we stop the iteration until the following criterion

$$\frac{\left\|x^{k+1}-x^k\right\|}{\|x^k\|} < 1e-3(\text{ for multichannel case } \frac{\left\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^k\right\|}{\|\boldsymbol{x}^k\|} < 1e-3)$$

satisfies. We measure the quality of the restored images by the improvement of signal to noise ratio (ISNR)

$$\text{ISNR}(x^*) = 10\log_{10}\frac{\|\underline{x}-x^*\|}{\|\underline{x}-d\|},$$

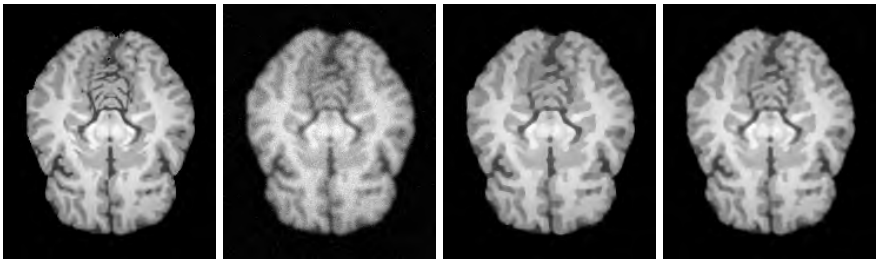
where $\underline{x}$ is the ground truth image, d is the observed image and x^* is the recovered image. For multichannel case, we have the similar definition of ISNR. For each model, the parameter α is tuned to obtain the highest ISNR. The performances of augmented Lagrangian methods are demonstrated in Figure 3–Figure 11.

(a) Original. Size: 512×512 (b) Blurry&Noisy (c) ALM for TV-L^2 ISNR:6.83dB t=1.60s

Fig. 3: Augmented Lagrangian method (ALM) for solving TV-L^2 model. (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',11,3)` and Gaussian noise with variation $1e-2$; (c) is the recovered result.

Figure 3 shows the results of augmented Lagrangian method for solving TV-L^2 model. In this experiment, we corrupt the clean image (size 512×512) with Gaussian blur and Gaussian noise. We set the parameters by following the recommendations in Wu and Tai (2010) and let $\beta = 10$. We report the recovered image and its ISNR in Figure 3(c). We also record the used CPU time t when the algorithm terminates. We can see that augmented Lagrangian method can solve TV-L^2 model efficiently and obtain high quality recovered image.

Figure 4 shows the results of augmented Lagrangian method for solving TV-L^2 model with box constraint and the comparisons with TV-L^2 model. In this experiment, the degraded image (size 217×181) is corrupted with Gaussian blur and Gaussian noise. We set the parameters $\beta = \beta_y = 10$ and $\beta_z = 400$. We report the



(a) Original. Size: 217×181 | (b) Blurry&Noisy | (c) ALM for TV-L^2 ISNR:3.77dB t=0.39s | (d) ALM for TVBox ISNR:4.26dB t=0.78s

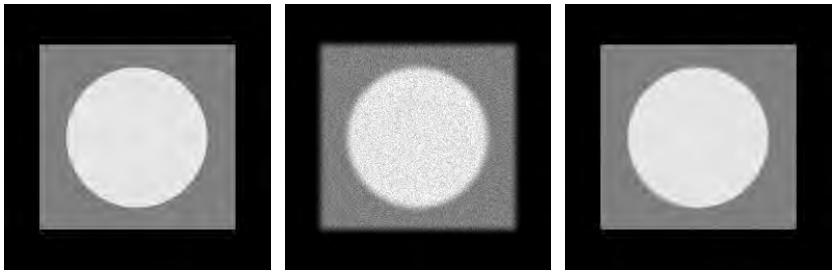
Fig. 4: Augmented Lagrangian method for solving TV-L^2 model with box constraint (TVBox). (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',5,1.5)` and Gaussian noise with variation $1e-3$; (c) and (d) are the recovered results.

recovered images and their ISNRs in Figure 4(c) and (d). We also record the used CPU times t when the algorithms terminate. We can see that augmented Lagrangian method can solve TV-L^2 model with box constraint efficiently and obtain high quality recovered image. The TV-L^2 model with box constraint gains higher ISNR than the TV-L^2 model.

(a) Original. Size: 512×512 | (b) Blurry&Noisy 50% salt & pepper | (c) ALM for TV-L^1 ISNR:20.34dB t=4.7s

Fig. 5: Augmented Lagrangian method for solving TV-L^1 model. (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',11,3)` and 50% salt & pepper noise; (c) is the recovered result.

Figure 5 and Figure 6 show the results of augmented Lagrangian methods for TV-L^1 model and TV-KL model. In the experiment for TV-L^1 model, the observed image (size 512×512) is degraded with Gaussian blur and 50% salt & pepper noise. We set $\beta_y = 20$ and $\beta_z = 100$. In the experiment for TV-KL model, the observed



(a) Original. Size: 256×256 | (b) Blurry&Noisy Poisson | (c) ALM for TV-KL ISNR:14.89dB t=2.5s

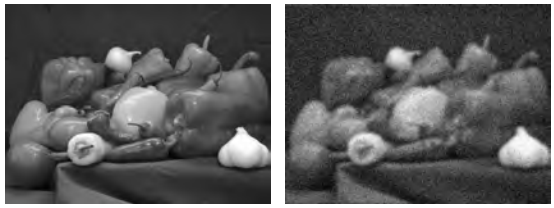
Fig. 6: Augmented Lagrangian method for solving TV-KL model. (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',11,3)` and Poisson noise; (c) is the recovered result.

image (size 256×256) is corrupted with Gaussian kernel and Poisson noise. We let $\beta_y = 20$ and $\beta_z = 20$. We can see that augmented Lagrangian methods can recover high quality images in these two experiments and the CPU costs are low.

(a) Lena in color. Size: $512 \times 512 \times 3$ | (b) Blurry&Noisy | (c) ALM for MTV ISNR:2.99dB t=6.21s

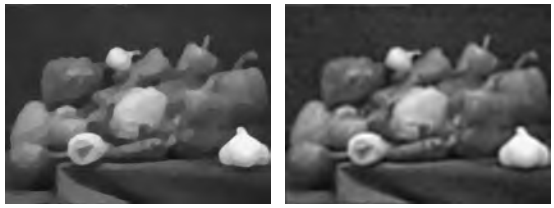
Fig. 7: Augmented Lagrangian method for multichannel TV (MTV) restoration. (b) is a corruption of (a) with within-channel Gaussian blur `fspecial('gaussian',21,5)` and Gaussian noise with variation $1e-3$; (c) is the recovered result.

Figure 7 shows the results of augmented Lagrangian method for multichannel TV restoration. In this experiment, the degraded image is generated by first blurring the ground truth image (size $512 \times 512 \times 3$) with within-channel Gaussian blur and then adding Gaussian noise to the blurred image. We set $\beta = 100$. We also can see that augmented Lagrangian method can restore high quality multichannel image with a low CPU cost.



(a) Original.
Size: 384×512

(b) Blurry&Noisy



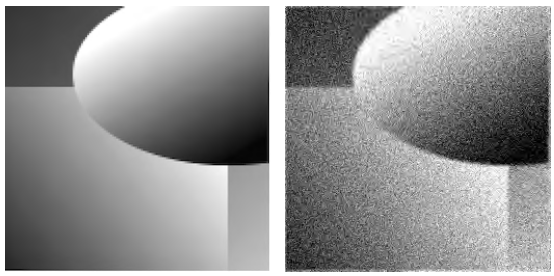
(c) ALM for TV-L^2
ISNR:8.87dB t=1.08s

(d) ALM for LLT
ISNR:8.99dB t=2.26s

Fig. 8: Augmented Lagrangian method for solving LLT model. (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',11,3)` and Gaussian noise with variation $1e-2$; (c) and (d) are the recovered results.

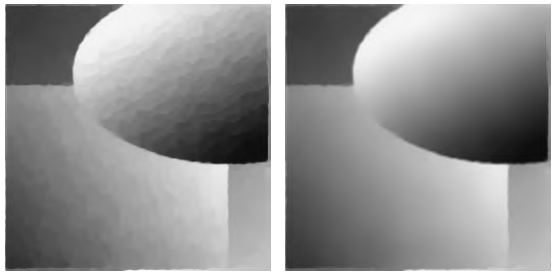
Figure 8 and Figure 9 show the results of augmented Lagrangian methods for solving LLT model and TGV model and the comparisons with TV-L^2 model. In the experiment for LLT model, the degraded image (size 384×512) is generated with Gaussian blur and Gaussian noise. We set $\beta = 10$. In the experiment for TGV model, the degraded image (size 256×256) is also generated with Gaussian blur and Gaussian noise. We let $(\alpha_0, \alpha_1) = (1.0, 0.1)$, $\beta_y = 10$ and $\beta_z = 20$. We report the recovered images and their ISNRs in Figure 8(c)-(d) and Figure 9(c)-(d). We also record the used CPU times t when the algorithms terminate. We can see that augmented Lagrangian method can solve LLT model and TGV model efficiently and obtain high quality recovered images. The LLT model and TGV model, which use high order regularization, can suppress the staircase effect well.

Figure 10 and Figure 11 show the results of augmented Lagrangian methods for solving Euler's elastica based image denoising model and mean curvature based image denoising model. Both these two models include curvature term in the regularization and are non-convex and highly nonlinear. We generate the degraded images Figure 10(b) and Figure 11(b) by adding Gaussian noise to the clean images Figure 10(a) and Figure 11(a) respectively. In the experiment for Euler's elastica based model, we use $\beta_y = 200$, $\beta_n = 500$ and $\beta_m = 1$. In the experiment for mean curvature based model, we use $\beta_y = 40$, $\beta_q = 1e5$, $\beta_n = 1e5$ and $\beta_m = 40$. We report the recovered images and their ISNRs and show the CPU costs in Figure 10(c) and



(a) Original.
Size: 256×256

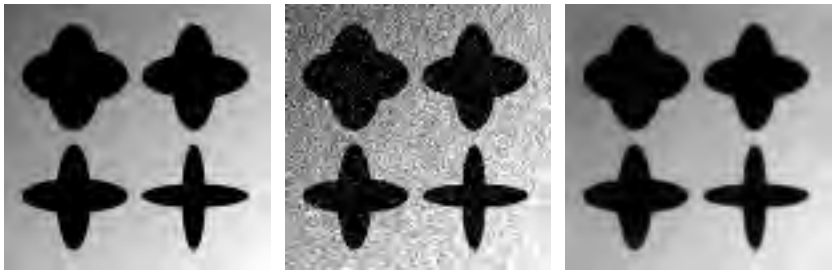
(b) Blurry&Noisy



(c) ALM for TV-L^2
ISNR:6.68dB t=0.40s

(d) ALM for TGV
ISNR:6.76dB t=1.84s

Fig. 9: Augmented Lagrangian method for solving TGV model. (b) is a corruption of (a) with Gaussian blur `fspecial('gaussian',5,1.5)` and Gaussian noise with variation $1e-2$; (c) and (d) are the recovered results.

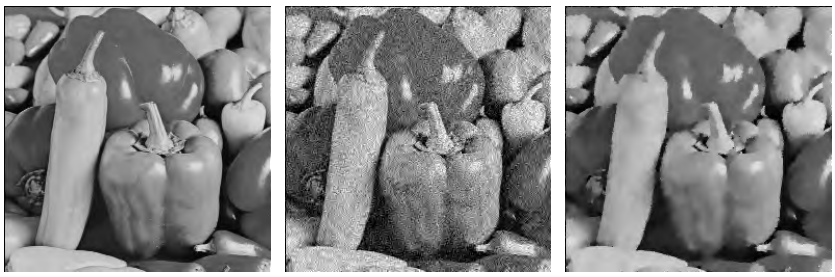


(a) Original.
Size: 100×100

(b) Noisy

(c) ALM for EE
ISNR:9.37dB t=1.52s

Fig. 10: Augmented Lagrangian method for solving Euler's elastica (EE) based image denoising model. (b) is a corruption of (a) with Gaussian noise with variation $1e-2$; (c) is the recovered result.



(a) Original. Size: 256×256 (b) Noisy (c) ALM for MC ISNR:8.10dB t=4.17s

Fig. 11: Augmented Lagrangian method for solving mean curvature (MC) based image denoising model. (b) is a corruption of (a) with Gaussian noise with variation $1e-2$; (c) is the recovered result.

Figure 11(c). We can see that augmented Lagrangian methods can solve non-convex curvature based models efficiently and obtain high quality recovered images.

Conclusions

In this survey, we have reviewed variable splitting and augmented Lagrangian methods for total variation related image restoration models. Due to the closed form solutions of subproblems and fast linear solvers like the FFT implementations, these methods are efficient for both total variation related convex models and non-convex Euler's elastica and mean curvature based models.

References

Acar R, Vogel CR (1994) Analysis of bounded variation penalty methods for ill-posed problems. Inverse Problems 10(6):1217–1229

Afonso MV, Bioucas-Dias JM, Figueiredo MAT (2010) Fast image recovery using variable splitting and constrained optimization. IEEE Trans Image Process 19(9):2345–2356

Afonso MV, Bioucas-Dias JM, Figueiredo MAT (2011) An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. IEEE Trans Image Process 20(3):681–695

Allison MJ, Ramani S, Fessler JA (2013) Accelerated regularized estimation of MR coil sensitivities using augmented Lagrangian methods. IEEE Trans Med Imaging 32(3):556–564

Aubert G, Aujol JF (2008) A variational approach to removing multiplicative noise. SIAM J Appl Math 68(4):925–946

Aubert G, Kornprobst P (2010) Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, 2nd edn. Springer, New York, NY

Aujol JF, Chambolle A (2005) Dual norms and image decomposition models. Int J Comput Vis 63(1):85–104

Aujol JF, Aubert G, Blanc-Féraud L, Chambolle A (2005) Image decomposition into a bounded variation component and an oscillating component. J Math Imaging Vis 22(1):71–88

Bae E, Shi J, Tai XC (2010) Graph cuts for curvature based image denoising. IEEE Trans Image Process 20(5):1199–1210

Bae E, Tai XC, Zhu W (2017) Augmented Lagrangian method for an Euler's elastica based segmentation model that promotes convex contours. Inverse Probl Imaging 11(1):1–23

Beck A, Teboulle M (2009) Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Trans Image Process 18(11):2419–2434

Bertalmio M, Cheng L, Osher S, Sapiro G (2001) Variational problems and partial differential equations on implicit surfaces. J Comput Phys 174(2):759–780

Bertalmio M, Vese L, Sapiro G, Osher S (2003) Simultaneous structure and texture image inpainting. IEEE Trans Image Process 12(8):882–889

Bertsekas DP (1996(first published 1982)) Constrained Optimization and Lagrange Multiplier Methods. Optimization and Neural Computation Series, Athena Scientific, Belmont, Mass

Bian W, Chen X (2015) Linearly constrained non-Lipschitz optimization for image restoration. SIAM J Imaging Sci 8(4):2294–2322

Blake A, Zisserman A (1987) Visual reconstruction. MIT Press, Cambridge, MA, USA

Blomgren P, Chan TF (1998) Color TV: Total variation methods for restoration of vector-valued images. IEEE Trans Image Process 7(3):304–309

Blomgren P, Chan TF, Mulet P, Wong CK (1997) Total variation image restoration: Numerical methods and extensions. In: In Proc. IEEE ICIP, pp 384–387

Boţ RI, Nguyen DK (2020) The proximal alternating direction method of multipliers in the nonconvex setting: Convergence analysis and rates. Math Methods Oper Res 45(2):682–712

Boyd S (2010) Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn 3(1):1–122

Bredies K, Holler M (2020) Higher-order total variation approaches and generalisations. Inverse Problems URL http://iopscience.iop.org/10.1088/1361-6420/ab8f80

Bredies K, Valkonen T (2011) Inverse problems with second-order total generalized variation constraints. In: Proceedings of SampTA, 2011, p 4

Bredies K, Kunisch K, Pock T (2010) Total generalized variation. SIAM J Imaging Sci 3(3):492–526

Bredies K, Pock T, Wirth B (2015) A convex, lower semicontinuous approximation of Euler's elastica energy. SIAM J Math Anal 47(1):566–613

Bresson X, Chan TF (2008) Fast dual minimization of the vectorial total variation norm and applications to color image processing. Inverse Probl Imaging 2(4):455–484

Brune C, Sawatzky A, Burger M (2009) Bregman-em-tv methods with application to optical nanoscopy. In: Tai XC, Mørken K, Lysaker M, Lie KA (eds) Scale Space and Variational Methods in Computer Vision, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 235–246

Cai X, Han D, Xu L (2013) An improved first-order primal-dual algorithm with a new correction step. J Global Optim 57(4):1419–1428

Caselles V, Chambolle A, Novaga M (2007) The discontinuity set of solutions of the TV denoising problem and some extensions. Multiscale Model Simul 6(3):879–894

Chambolle A (2004) An algorithm for total variation minimization and applications. J Math Imaging Vis 20(1/2):89–97

Chambolle A, Lions PL (1997) Image recovery via total variation minimization and related problems. Numer Math 76(2):167–188

Chambolle A, Pock T (2011) A first-order primal-dual algorithm for convex problems with applications to imaging. J Math Imaging Vis 40(1):120–145

Chambolle A, Pock T (2019) Total roto-translational variation. Numer Math 142(3):611–666

Chan R, Lanza A, Morigi S, Sgallari F (2018) Convex non-convex image segmentation. Numer Math 138(3):635–680

Chan RH, Ma J (2012) A multiplicative iterative algorithm for box-constrained penalized likelihood image restoration. IEEE Trans Image Process 21(7):3168–3181

Chan RH, Tao M, Yuan X (2013) Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers. SIAM J Imaging Sci 6(1):680–697

Chan SH, Khoshabeh R, Gibson KB, Gill PE, Nguyen TQ (2011) An augmented Lagrangian method for total variation video restoration. IEEE Trans Image Process 20(11):3097–3111

Chan T, Wong CK (1998) Total variation blind deconvolution. IEEE Trans Image Process 7(3):370–375

Chan T, Golub G, Mulet P (1999) A nonlinear primal-dual method for total variation-based image restoration. SIAM J Sci Comput 20(6):1964–1977

Chan T, Marquina A, Mulet P (2000) High-order total variation-based image restoration. SIAM J Sci Comput 22(2):503–516

Chan TF, Mulet P (1999) On the convergence of the lagged diffusivity fixed point method in total variation image restoration. SIAM J Numer Anal 36(2):354–367

Chan TF, Shen J (2005) Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods. Society for Industrial and Applied Mathematics, Philadelphia

Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277

Chan TF, Kang SH, Shen J (2002) Euler's elastica and curvature-based inpainting. SIAM J Appl Math 63(2):564–592

Chan TF, Yip AM, Park FE (2005) Simultaneous total variation image inpainting and blind deconvolution. Int J Imag Syst Tech 15(1):92–102

Chan TF, Esedoglu S, Nikolova M (2006) Algorithms for finding global minimizers of image segmentation and denoising models. SIAM J Appl Math 66(5):1632–1648

Chang H, Lou Y, Ng M, Zeng T (2016) Phase retrieval from incomplete magnitude information via total variation regularization. SIAM J Sci Comput 38(6):A3672–A3695

Chang H, Lou Y, Duan Y, Marchesini S (2018) Total variation–based phase retrieval for poisson noise removal. SIAM J Imaging Sci 11(1):24–55

Chartrand R, Wohlberg B (2013) A nonconvex ADMM algorithm for group sparsity with sparse groups. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp 6009–6013

Chen C, Chan RH, Ma S, Yang J (2015) Inertial proximal admm for linearly constrained separable convex optimization. SIAM J Imaging Sci 8(4):2239–2267

Chen C, Chen Y, Ouyang Y, Pasiliao E (2018) Stochastic accelerated alternating direction method of multipliers with importance sampling. J Optim Theory Appl 179(2):676–695

Chen DQ, Cheng LZ, Su F (2012a) A new TV-Stokes model with augmented Lagrangian method for image denoising and deconvolution. J Sci Comput 51(3):505–526

Chen F, Shen L, Xu Y, Zeng X (2014) The moreau envelope approach for the l1/tv image denoising model. Inverse Probl Imaging 8(1):53–77

Chen T, Wotao Yin, Xiang Sean Zhou, Comaniciu D, Huang T (2006a) Total variation models for variable lighting face recognition. IEEE Trans Pattern Anal Mach Intell 28(9):1519–1524

Chen X, Xu F, Ye Y (2010) Lower bound theory of nonzero entries in solutions of ℓ_2-ℓ_p minimization. SIAM J Imaging Sci 32(5):2832–2852

Chen X, Ng MK, Zhang C (2012b) Non-Lipschitz ℓ_p-regularization and box constrained model for image restoration. IEEE Trans Image Process 21(12):4709–4721

Chen Y, Levine S, Rao M (2006b) Variable exponent, linear growth functionals in image restoration. SIAM J Appl Math 66(4):1383–1406

Chen Z, Jin X, Li L, Wang G (2013) A limited-angle CT reconstruction method based on anisotropic TV minimization. Physics in Medicine and Biology 58(7):2119–2141

Coll B, Duran J, Sbert C (2015) Half-linear regularization for nonconvex image restoration models. Inverse Probl Imaging 9(2):337–370

Cui Y, Li X, Sun D, Toh Kc (2016) On the convergence properties of a majorized alternating direction method of multipliers for linearly constrained convex optimization problems with coupled objective functions. J Optim Theory Appl 169(3):1013–1041

Deng LJ, Glowinski R, Tai XC (2019) A new operator splitting method for the euler elastica model for image smoothing. SIAM J Imaging Sci 12(2):1190–1230

Deng W, Yin W (2016) On the global and linear convergence of the generalized alternating direction method of multipliers. J Sci Comput 66(3):889–916

Dong B, Zhang Y (2013) An efficient algorithm for l0 minimization in wavelet frame based image restoration. J Sci Comput 54(2-3):350–368

Dong F, Zhang H, Kong D (2012) Nonlocal total variation models for multiplicative noise removal using split bregman iteration. Mathematical and Computer Modelling 55(3):939–954

Dong Y, Zeng T (2013) A convex variational model for restoring blurred images with multiplicative noise. SIAM J Imaging Sci 6(3):1598–1625

Dong Y, Hintermuller M, Neri M (2009) An efficient primal-dual method for l^1tv image restoration. SIAM J Imaging Sci 2(4):1168–1189

Donoho DL (1995) De-noising by soft-thresholding. IEEE Trans Inform Theory 41(3):613–627

Duan Y, Wang Y, Hahn J (2013) A fast augmented Lagrangian method for Euler's elastica models. Numer Math Theory Methods Appl 006(001):47–71

Duan Y, Huang W, Zhou J, Chang H, Zeng T (2014) A two-stage image segmentation method using Euler's elastica regularized Mumford-Shah model. In: International Conference on Pattern Recognition, pp 118–123

Ekeland I, Témam R (1999) Convex Analysis and Variational Problems. SIAM

Esedoglu S, March R (2003) Segmentation with depth but without detecting junctions. J Math Imaging Vis 18(1):7–15

Esser E, Zhang X, Chan TF (2010) A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. SIAM J Imaging Sci 3(4):1015–1046

Fang C, Cheng F, Lin Z (2017) Faster and non-ergodic o(1/k) stochastic alternating direction method of multipliers. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds) Advances in Neural Information Processing Systems 30, Curran Associates, Inc., pp 4476–4485

Fazel M, Pong TK, Sun D, Tseng P (2013) Hankel matrix rank minimization with applications to system identification and realization. SIAM J Matrix Anal & Appl 34(3):946–977

Feng X, Wu C, Zeng C (2018) On the local and global minimizers of ℓ_0 gradient regularized model with box constraints for image restoration. Inverse Problems 34(9):095,007

Freiberger M, Clason C, Scharfetter H (2010) Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach. Appl Opt, AO 49(19):3741–3747

Gao Y, Wu C (2020) On a general smoothly truncated regularization for variational piecewise constant image restoration: Construction and convergent algorithms. Inverse Problems 36(4):045,007

Gao Y, Liu F, Yang X (2018) Total generalized variation restoration with non-quadratic fidelity. Multidim Syst Sign Process 29(4):1459–1484

Glowinski R (1984) Numerical Methods for Nonlinear Variational Problems. Scientific Computation, Springer Berlin Heidelberg, Berlin, Heidelberg

Glowinski R (2014) On alternating direction methods of multipliers: A historical perspective. In: Fitzgibbon W, Kuznetsov YA, Neittaanmäki P, Pironneau O (eds) Modeling, Simulation and Optimization for Science and Technology, Springer Netherlands, Dordrecht, pp 59–82

Glowinski R (2015) Variational Methods for the Numerical Solution of Nonlinear Elliptic Problems. Society for Industrial and Applied Mathematics, Philadelphia, PA

Glowinski R, Tallec PL (1989) Augmented Lagrangians and Operator-Splitting Methods in Nonlinear Mechanics. SIAM, Philadelphia

Glowinski R, Osher SJ, Yin W (eds) (2016a) Splitting Methods in Communication, Imaging, Science, and Engineering. Springer, Cham

Glowinski R, Pan TW, Tai XC (2016b) Some facts about operator-splitting and alternating direction methods. In: Glowinski R, Osher SJ, Yin W (eds) Splitting Methods in Communication, Imaging, Science, and Engineering, Springer International Publishing, Cham, pp 19–94

Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. SIAM J Imaging Sci 2(2):323–343

Guo W, Huang F (2009) Adaptive total variation based filtering for MRI images with spatially inhomogeneous noise and artifacts. In: International Symposium on Biomedical Imaging, pp 101–104

Guo X, Li F, Ng MK (2009) A fast $\ell 1$-tv algorithm for image restoration. SIAM J Sci Comput 31(3):2322–2341

Güven HE, Güngör A, Çetin M (2016) An augmented Lagrangian method for complex-valued compressed SAR imaging. IEEE Trans Comput Imag 2(3):235–250

Hahn J, Wu C, Tai XC (2012) Augmented Lagrangian method for generalized TV-Stokes model. J Sci Comput 50(2):235–264

Han D, Yuan X, Zhang W (2014) An augmented Lagrangian based parallel splitting method for separable convex minimization with applications to image processing. Math Comp 83(289):2263–2291

He B, Yuan X (2012) On the $o(1/n)$ convergence rate of the douglas-rachford alternating direction method. SIAM J Numer Anal 50(2):700–709

He B, Yuan X (2015) On non-ergodic convergence rate of Douglas–Rachford alternating direction method of multipliers. Numer Math 130(3):567–577

He B, Liao LZ, Han D, Yang H (2002) A new inexact alternating directions method for monotone variational inequalities. Math Program, Series B 92(1):103–118

He B, You Y, Yuan X (2014) On the convergence of primal-dual hybrid gradient algorithm. SIAM J Imaging Sci 7(4):2526–2537

He B, Ma F, Yuan X (2016) Convergence study on the symmetric version of admm with larger step sizes. SIAM J Imaging Sci 9(3):1467–1501

He B, Ma F, Yuan X (2020) Optimally linearizing the alternating direction method of multipliers for convex programming. Comput Optim Appl 75(2):361–388

Hestenes MR (1969) Multiplier and gradient methods. J Optim Theory Appl 4(5):303–320

Hintermüller M, Wu T (2013) Nonconvex TV^q-models in image restoration: Analysis and a trust-region regularization–based superlinearly convergent solver. SIAM J Imaging Sci 6(3):1385–1415

Hong M, Luo ZQ, Razaviyayn M (2016) Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. SIAM J Optim 26(1):337–364

Houska B, Frasch J, Diehl M (2016) An augmented Lagrangian based algorithm for distributed nonconvex optimization. SIAM J Optim 26(2):1101–1127

Huang Y, Ng MK, Wen YW (2008) A fast total variation minimization method for image restoration. Multiscale Model Simul 7(2):774–795

Ilbey S, Top CB, Çukur T, Saritas EU, Güven HE (2017) Image reconstruction for magnetic particle imaging using an augmented lagrangian method. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp 64–67

Jia X, Lou Y, Lewis J, Li R, Gu X, Men C, Song W, Jiang SB (2011) Gpu-based fast low-dose cone beam CT reconstruction via total variation. Journal of X-ray Science and Technology 19(2):139–154

Jiao Y, Jin Q, Lu X, Wang W (2016) Alternating direction method of multipliers for linear inverse problems. SIAM J Numer Anal 54(4):2114–2137

Jiao Y, Jin Q, Lu X, Wang W (2017) Preconditioned alternating direction method of multipliers for inverse problems with constraints. Inverse Problems 33(2):025,004

Jin Z, Yang X (2011) A variational model to remove the multiplicative noise in ultrasound images. J Math Imaging Vis 39(1):62–74

Kadkhodaie M, Christakopoulou K, Sanjabi M, Banerjee A (2015) Accelerated alternating direction method of multipliers. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA, KDD 15, p 497506

Kang SH, Zhu W, Jianhong J (2014) Illusory shapes via corner fusion. SIAM J Imaging Sci 7(4):1907–1936

Koko J, Jehan-Besson S (2010) An augmented Lagrangian method for TVg +L1-norm minimization. J Math Imaging Vis 38(3):182–196
Lai R, Chan TF (2011) A framework for intrinsic image processing on surfaces. Comput Vis Image Und 115(12):1647–1661
Lanza A, Morigi S, Sgallari F (2016) Convex image denoising via non-convex regularization with parameter selection. J Math Imaging Vis 56(2):195–220
Le T, Chartrand R, Asaki TJ (2007) A variational approach to reconstructing images corrupted by Poisson noise. J Math Imaging Vis 27:257–263
Le TM, Vese LA (2005) Image decomposition using total variation and div(bmo). Multiscale Model Simul 4(2):390–423
Lellmann J, Strekalovskiy E, Koetter S, Cremers D (2013) Total variation regularization for functions with values in a manifold. In: The IEEE International Conference on Computer Vision (ICCV), pp 2944–2951
Li C, Yin W, Jiang H, Zhang Y (2013) An efficient augmented Lagrangian method with applications to total variation minimization. Comput Optim Appl 56(3):507–530
Li F, Li Z, Pi L (2010) Variable exponent functionals in image restoration. Adv Comput Math 216(3):870882
Li G, Pong TK (2015) Global convergence of splitting methods for nonconvex composite optimization. SIAM J Optim 25(4):2434–2460
Li H, Lin Z (2019) Accelerated alternating direction method of multipliers: An optimal o(1/k) nonergodic analysis. J Sci Comput 79(2):671–699
Li X, Sun D, Toh KC (2016) A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions. Math Program 155(1-2):333–373
Liu J, Zheng X (2017) A block nonlocal tv method for image restoration. SIAM J Imaging Sci 10(2):920–941
Liu X (2016) Augmented Lagrangian method for total generalized variation based Poissonian image restoration. Comput Math Appl 71(8):1694–1705
Liu Z, Wali S, Duan Y, Chang H, Wu C, Tai XC (2018) Proximal ADMM for Euler's elastica based image decomposition model. Numer Math Theory Methods Appl 12(2):370–402
Liu Z, Lai R, Zhang H, Wu C (2019) Triangulated surface denoising using high order regularization with dynamic weights. SIAM J Sci Comput 41(1):B1–B26
Lou Y, Zhang X, Osher S, Bertozzi AL (2010) Image recovery via nonlocal operators. J Sci Comput 42(2):185–197
Lysaker M, Lundervold A, Tai X (2003) Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. IEEE Trans Image Process 12(12):1579–1590
Ma L, Yu J, Zeng T (2013) Sparse representation prior and total variation–based image deblurring under impulse noise. SIAM J Imaging Sci 6(4):2258–2284
Marquina A, Osher SJ (2008) Image super-resolution by TV-regularization and Bregman iteration. J Sci Comput 37(3):367–382
Masnou S, Morel J (2005) On a variational theory of image amodal completion. Rendiconti Del Seminario Matematico Della Universita? Di Padova 116(4):211–252
Masnou S, Morel JM (1998) Level lines based disocclusion. In: Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269), pp 259–263 vol.3
Micchelli CA, Shen L, Xu Y (2011) Proximity algorithms for image models: denoising. Inverse Problems 27(4):045,009
Micchelli CA, Shen L, Xu Y, Zeng X (2013) Proximity algorithms for the l1/tv image denoising model. Adv Comput Math 38(2):401–426
Myllykoski M, Glowinski R, Karkkainen T, Rossi T (2015) A new augmented Lagrangian approach for L^1-mean curvature image denoising. SIAM J Imaging Sci 8(1):95–125
Ng MK, Weiss P, Yuan X (2010) Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. SIAM J Sci Comput 32(5):2710–2736

Ng MK, Xiaoming Yuan, Wenxing Zhang (2013) Coupled variational image decomposition and restoration model for blurred cartoon-plus-texture images with missing pixels. IEEE Trans Image Process 22(6):2233–2246

Nien H, Fessler JA (2015) Fast x-ray ct image reconstruction using a linearized augmented Lagrangian method with ordered subsets. IEEE Trans Med Imaging 34(2):388–399

Nikolova M (2002) Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers. SIAM J Numer Anal 40(3):965–994

Nikolova M (2004) A variational approach to remove outliers and impulse noise. J Math Imaging Vis 20(1-2):99–120

Nikolova M (2005) Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. Multiscale Model Simul 4(3):960–991

Nikolova M, Ng MK, Zhang S, Ching WK (2008) Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. SIAM J Imaging Sci 1(1):2–25

Nikolova M, Ng MK, Chi-Pan Tam (2010) Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction. IEEE Trans Image Process 19(12):3073–3088

Osher S, Shi Z, Zhu W (2017) Low dimensional manifold model for image processing. SIAM J Imaging Sci 10(4):1669–1690

Ouyang H, He N, Tran LQ, Gray A (2013) Stochastic alternating direction method of multipliers. In: Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28, JMLR.org, ICML13, p I8088

Ouyang Y, Chen Y, Lan G, Pasiliao E (2015) An accelerated linearized alternating direction method of multipliers. SIAM J Imaging Sci 8(1):644–681

Pang ZF, Meng G, Li H, Chen K (2020) Image restoration via the adaptive TVp regularization. Comput Math Appl 80(5):569–587

Persson M, Bone D, Elmqvist H (2001) Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. Physics in Medicine and Biology 46(3):853–866

Peyré G, Bougleux S, Cohen L (2008) Non-local regularization of inverse problems. In: Proceedings of the 10th European Conference on Computer Vision: Part III, Springer-Verlag, Berlin, Heidelberg, ECCV 08, p 5768

Powell MJD (1969) A method for nonliner constraints in minimization problem. In: Fletcher R (ed) Optimization, AcademicPres, New York, pp 283–298

Ramani S, Fessler JA (2011) Parallel MR image reconstruction using augmented Lagrangian methods. IEEE Trans Med Imaging 30(3):694–706

Ritschl L, Bergner F, Fleischmann C, Kachelrieß M (2011) Improved total variation-based CT image reconstruction applied to clinical data. Physics in Medicine and Biology 56(6):1545–1561

Rockafellar RT (1974) Augmented Lagrange multiplier functions and duality in nonconvex programming. SIAM Journal on Control 12(2):268–285

Rockafellar RT (1976) Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math Methods Oper Res 1(2):97–116

Rockafellar RT, Wets RJB (1998) Variational Analysis. Springer, Berlin, Heidelberg

Rudin L, Osher S (1994) Total variation based image restoration with free local constraints. In: Proceedings of 1st International Conference on Image Processing, IEEE Comput. Soc. Press, Austin, TX, USA, vol 1, pp 31–35

Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D 60:259–268

Rudin L, Lions PL, Osher S (2003) Multiplicative denoising and deblurring: Theory and algorithms. In: Geometric Level Set Methods in Imaging, Vision, and Graphics, Springer New York, New York, NY, pp 103–119

Sapiro G, Ringach D (1996) Anisotropic diffusion of multivalued images with applications to color filtering. IEEE Trans Image Process 5:1582–1586

Sciacchitano F, Dong Y, Zeng T (2015) Variational approach for restoring blurred images with cauchy noise. SIAM J Imaging Sci 8(3):1894–1922

Selesnick I, Lanza A, Morigi S, Sgallari F (2020) Non-convex total variation regularization for convex denoising of signals. J Math Imaging Vis 62(6):825–841

Setzer S, Steidl G, Teuber T (2010) Deblurring poissonian images by split bregman techniques. J Vis Commun Image Represent 21:193–199

Shen J, Chan T (2002) Mathematical models for local nontexture inpaintings. SIAM J Appl Math 62(3):1019–1043

Shirai K, Okuda M (2014) FFT based solution for multivariable L2 equations using KKT system via FFT and efficient pixel-wise inverse calculation. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 2629–2633

Sidky EY, Pan X (2008) Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Physics in Medicine and Biology 53(17):4777–4807

Sidky EY, Duchin Y, Pan X, Ullberg C (2011) A constrained, total-variation minimization algorithm for low-intensity x-ray ct. Medical Physics 38(S1):S117–S125

Strong D, Chan T (2003) Edge-preserving and scale-dependent properties of total variation regularization. Inverse Problems 19(6):S165–S187

Sun L, Chen K (2014) A new iterative algorithm for mean curvature-based variational image denoising. BIT Numer Math 54(2):523–553

Tai XC, Wu C (2009) Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model. In: Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings, pp 502–513

Tai XC, Hahn J, Chung GJ (2011) A fast algorithm for Euler's elastica model using augmented Lagrangian method. SIAM J Imaging Sci 4(1):313–344

Tan L, Li L, Liu W, Sun J, Zhang M (2020) A novel Euler's elastica-based segmentation approach for noisy images using the progressive hedging algorithm. J Math Imaging Vis 62:98–119

Themelis A, Patrinos P (2020) Douglas–Rachford splitting and ADMM for nonconvex optimization: Tight convergence results. SIAM J Optim 30(1):149–181

Tian Z, Jia X, Yuan K, Pan T, Jiang SB (2011) Low dose CT reconstruction via edge-preserving total variation regularization. Physics in Medicine and Biology 56(18):5949–5967

Vese LA, Osher SJ (2003) Modeling textures with total variation minimization and oscillating patterns in image processing. J Sci Comput 19(1/3):553–572

Vese LA, Osher SJ (2004) Image denoising and decomposition with total variation minimization and oscillatory functions. J Math Imaging Vis 20(1/2):7–18

Vogel C, Oman M (1998) Fast, robust total variation-based reconstruction of noisy, blurred images. IEEE Trans Image Process 7(6):813–824

Wahlberg B, Boyd S, Annergren M, Wang Y (2012) An admm algorithm for a class of total variation regularized estimation problems. IFAC Proceedings Volumes 45(16):83–88

Wang F, Cao W, Xu Z (2018) Convergence of multi-block Bregman ADMM for nonconvex composite problems. Science China Information Sciences 61(12):122,101

Wang X, Yuan X (2012) The linearized alternating direction method of multipliers for dantzig selector. SIAM J Sci Comput 34(5):A2792–A2811

Wang Y, Yang J, Yin W, Zhang Y (2008) A new alternating minimization algorithm for total variation image reconstruction. SIAM J Imaging Sci 1(3):248–272

Wang Y, Yin W, Zeng J (2019) Global convergence of ADMM in nonconvex nonsmooth optimization. J Sci Comput 78(1):29–63

Weinmann A, Demaret L, Storath M (2014) Total variation regularization for manifold-valued data. SIAM J Imaging Sci 7(4):2226–2257

Wen Y, Chan RH, Zeng T (2016) Primal-dual algorithms for total variation based image restoration under poisson noise. Sci China Math 59(1):141–160

Wu C, Tai X (2010) Augmented Lagrangian method, dual methods, and split bregman iteration for ROF, vectorial TV, and high order models. SIAM J Imaging Sci 3(3):300–339

Wu C, Zhang J, Tai XC (2011) Augmented Lagrangian method for total variation restoration with non-quadratic fidelity. Inverse Probl Imaging 5(1):237–261

Wu C, Zhang J, Duan Y, Tai XC (2012) Augmented lagrangian method for total variation based image restoration and segmentation over triangulated surfaces. J Sci Comput 50(1):145–166

Wu C, Liu Z, Wen S (2018) A general truncated regularization framework for contrast-preserving variational signal and image restoration: Motivation and implementation. Sci China Math 61(9):1711–1732

Wu X, Zheng J, Wu C, Cai Y (2013) Variational structure-texture image decomposition on manifolds. Signal Processing 93(7):1773–1784

Xiao YH, Song HN (2012) An inexact alternating directions algorithm for constrained total variation regularized compressive sensing problems. J Math Imaging Vis 44(2):114–127

Xu J, Zhao Y, Li H, Zhang P (2019) An image reconstruction model regularized by edge-preserving diffusion and smoothing for limited-angle computed tomography. Inverse Problems 35(8):085,004

Xu L, Lu C, Xu Y, Jia J (2011) Image smoothing via l_0 gradient minimization. ACM Trans Graph 30(6):112

Yan M, Duan Y (2020) Nonlocal elastica model for sparse reconstruction. J Math Imaging Vis 62:532–548

Yan M, Yin W (2016) Self equivalence of the alternating direction method of multipliers. In: Glowinski R, Osher SJ, Yin W (eds) Splitting Methods in Communication, Imaging, Science, and Engineering, Springer International Publishing, Cham, pp 165–194

Yang J, Yin W, Zhang Y, Wang Y (2009a) A fast algorithm for edge-preserving variational multichannel image restoration. SIAM J Imaging Sci 2(2):569–592

Yang J, Zhang Y, Yin W (2009b) An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. SIAM J Sci Comput 31(4):2842–2865

Yang J, Yu H, Jiang M, Wang G (2010) High-order total variation minimization for interior tomography. Inverse Problems 26(3):035,013

Yang L, Pong TK, Chen X (2017) Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction. SIAM J Imaging Sci 10(1):74–110

Yang WH, Han D (2016) Linear convergence of the alternating direction method of multipliers for a class of convex optimization problems. SIAM J Numer Anal 54(2):625–640

Yashtini M, Kang SH (2016) A fast relaxed normal two split method and an effective weighted TV approach for Euler's elastica image inpainting. SIAM J Imaging Sci 9(4):1552–1581

You J, Jiao Y, Lu X, Zeng T (2019) A nonconvex model with minimax concave penalty for image restoration. J Sci Comput 78(2):1063–1086

Yue H, Yang Q, Wang X, Yuan X (2018) Implementing the alternating direction method of multipliers for big datasets: A case study of least absolute shrinkage and selection operator. SIAM J Sci Comput 40(5):A3121–A3156

Zeng C, Wu C (2018) On the edge recovery property of noncovex nonsmooth regularization in image restoration. SIAM J Numer Anal 56(2):1168–1182

Zeng C, Wu C (2019) On the discontinuity of images recovered by noncovex nonsmooth regularized isotropic models with box constraints. Adv Comput Math 45(2):589–610

Zeng C, Jia R, Wu C (2019a) An iterative support shrinking algorithm for non-Lipschitz optimization in image restoration. J Math Imaging Vis 61(1):122–139

Zeng C, Wu C, Jia R (2019b) Non-Lipschitz models for image restoration with impulse noise removal. SIAM J Imaging Sci 12(1):420–458

Zhang H, Wu C, Zhang J, Deng J (2015) Variational mesh denoising using total variation and piecewise constant function space. IEEE Trans Vis Comput Graphics 21(7):873–886

Zhang H, Wang L, Duan Y, Li L, Hu G, Yan B (2017a) Euler's elastica strategy for limited-angle computed tomography image reconstruction. IEEE Trans Nucl Sci 64(8):2395–2405

Zhang J (2018) A fast linearised Augmented lagrangian method for a mean curvature based model. East Asian J Appl Math 8(3):463–476

Zhang J, Wu C (2011) Fast optimization for multichannel total variation minimization with nonquadratic fidelity. Signal Processing 91(8):1933–1940

Zhang J, Chen R, Deng C, Wang S (2017b) Fast linearized augmented lagrangian method for euler's elastica model. Numer Math Theory Methods Appl 10(01):98–115

Zhao XL, Wang F, Ng MK (2014) A new convex optimization model for multiplicative noise and blur removal. SIAM J Imaging Sci 7(1):456–475
Zhu W, Chan T (2012) Image denoising using mean curvature of image surface. SIAM J Imaging Sci 5(1):1–32
Zhu W, Chan T, Esedo Lu S (2006) Segmentation with depth: A level set approach. SIAM J Sci Comput 28(5):1957–1973
Zhu W, Tai XC, Chan T (2013a) Augmented Lagrangian method for a mean curvature based image denoising model. Inverse Probl Imaging 7(4):1409–1432
Zhu W, Tai XC, Chan T (2013b) Image segmentation using euler's elastica as the regularization. J Sci Comput 57(2):414–438
Zhu Y, Zhao M, Zhao Y, Li H, Zhang P (2012) Noise reduction with low dose CT data based on a modified rof model. Opt Express 20(16):17,987–18,004