Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2017.m1653

Vol. **10**, No. 4, pp. 852-871
November 2017

Learning Non-Negativity Constrained Variation for Image Denoising and Deblurring

Tengda Wei[1], Linshan Wang[2], Ping Lin[3], Jialing Chen[3], Yangfan Wang[4,*]and Haiyong Zheng[5]

[1] *College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao 266100, China*
[2] *College of Mathematics, Ocean University of China, Qingdao 266100, China*
[3] *Department of Mathematics, University of Dundee, Dundee DD1 4HN, UK*
[4] *College of Marine Life Science, Ocean University of China, Qingdao 266100, China*
[5] *College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China*

Received 13 September 2016; Accepted (in revised version) 24 March 2017

Abstract. This paper presents a heuristic Learning-based Non-Negativity Constrained Variation (L-NNCV) aiming to search the coefficients of variational model automatically and make the variation adapt different images and problems by supervised-learning strategy. The model includes two terms: a problem-based term that is derived from the prior knowledge, and an image-driven regularization which is learned by some training samples. The model can be solved by classical ε-constraint method. Experimental results show that: the experimental effectiveness of each term in the regularization accords with the corresponding theoretical proof; the proposed method outperforms other PDE-based methods on image denoising and deblurring.

AMS subject classifications: 35A15, 35Q93
Key words: Learning idea, TV-based model, constraint, ε-constraint method, image restoration.

1. Introduction

Variational methods have been widely applied to various areas of image restoration and remain active in mathematical research of image processing. One of the most remarkable work is total variation (TV) which was first introduced into image denoising in the seminal work [36] by Rudin, Osher and Fatemi. The mostly used version of discrete TV [33, 38] is given by

$$\|u\|_{TV} = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1}\sqrt{(u_{i+1,j}-u_{i,j})^2+(u_{i,j+1}-u_{i,j})^2}, \tag{1.1}$$

*Corresponding author. *Email addresses:* tdwei123@163.com (T. Wei), wangls@ouc.edu.cn (L. Wang), plin@maths.dundee.ac.uk (P. Lin), j.z.chen@dundee.ac.uk (J. Chen), yfwang@ouc.edu.cn (Y. Wang), zhenghaiyong@ouc.edu.cn (H. Zheng)

 ©2017 Global-Science Press

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

which is the discretization of $\int_\Omega |\nabla u| d\boldsymbol{x}$ by pixel. The unconstrained version of TV-based model reads

$$\min_u \frac{1}{2}\|Ku - f\|^2 + \beta\|u\|_{TV}, \tag{1.2}$$

where $f : \Omega \to \mathbb{R}$ is the degraded image, $u : \Omega \to \mathbb{R}$ is the recovered image, K is the blurring operator, and the coefficient β usually depends on the noise level. The first term of (1.2) is the l_2 norm of $Ku - f$ which guarantees inheritance from the observed image f, and the second term is total variation which helps to remove the noise. This TV-based deblurring model solves different problems for different choices of K, for example, $K = I$ for image denoising. In this paper, we assume that K is derived from a point spread function (PSF), namely, K is known. Owing to excellent effectiveness of total variation on reducing noise and preserving edges, TV-based models have ignited plenty of research in dealing with image denoising and deblurring [12, 19, 21, 41, 46, 47], image inpainting [6, 10, 14, 15], image superresolution [13], image segmentation [11, 16, 26, 28, 32, 40] and other image processing problems [8].

Adding the non-negativity constraint to these TV-based models helps to recover the image when the constraint is physically meaningful. For example, images whose pixels represent the number of photon pairs must be non-negative. In contrast to clipping the solution to TV-based model, enforcing the constraint in the model can achieve better performance [17, 18]. Besides, the non-negative output images are the requirement of some applications such as the medical imaging [20, 34], gamma ray spectral analysis [31, 37] and so on. Krishnan et al. [25, 26] focused on the TV-based deblurring model with non-negativity constraint, called NNCGM, and solved it with the primal-dual active-set strategy [9, 24]. The interested reader is referred to [3, 22, 44] for other constrained algorithms.

All the above TV-based models are manually designed with the insight of individual problems, which sometimes restricts the strength of these TV-based models. Recently, a learning-based PDEs (L-PDE) model was proposed by Liu et al. in their seminal work [29]. Based on the proficiency of diffusion, they combined the differential operators with some image-driven coefficients to make PDEs adaptive to different problems. The model was proven to be successful in image denoising, edge detection and image segmentation [29].

Inspired by L-PDE methods, we propose a learning-based non-negativity constrained variation which extends the traditional TV-based models. The proposed variation contains two terms: a problem-based term that is derived from the prior knowledge, and an image-driven regularization which is learned by some training samples so as to make the variation adaptive to different images as well as problems. With the help of supervised-learning strategy, the optimal coefficients of the variation are searched automatically.

The rest of the paper is organized as follows. In Section 2, we propose a heuristic non-negativity constrained variation (NNCV) model, and then we give a learning-based form of our proposed model (L-NNCV) and solve it with ε-constraint method. Section 3 presents experimental results and Section 4 concludes this paper. Appendices A-B give supplementary materials containing extended proofs and mathematical derivations.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

2. Learning non-negativity constrained variation

2.1. NNCV model

In this subsection, we propose a class of non-negativity constrained variation (NNCV) model

$$\min_{u,u\geq 0} P(f,u)+V(\boldsymbol{\beta},u), \tag{2.1}$$

where $P(f,u)$ is a problem-based term derived from the prior knowledge, for example, $P(f,u)=\frac{1}{2}\|Ku-f\|^2$ for image deblurring while $P(f,u)=\langle s,u\rangle$ where $s=(f-c_1)^2-(f-c_2)^2$ for image segmentation, and $V(\boldsymbol{\beta},u):=\beta_1 v_1(u)+\frac{1}{2}\beta_2 v_2(u)+\frac{1}{2}\beta_3 v_3(u)$ is an image-driven regularization including three terms: $v_1(u)$ is the total variation that has been successfully applied to various areas due to its mathematical tractability; $v_2(u)=\|u\|^2$ makes the variation well conditioned; $v_3(u)=\|\nabla u\|^2$ is the strong smoothing term [43]. As controllers of the solution process, the coefficients of these three terms vary with iterations to make the variation adaptive to different images under different conditions. Obviously, TV-based deblurring model [36] and TV-based segmentation model [11] are special cases of our unconstrained NNCV model with fixed coefficients.

The NNCV model for image denoising and deblurring reads

$$\min_{u,u\geq 0} \frac{1}{2}\|Ku-f\|^2+V(\boldsymbol{\beta},u). \tag{2.2}$$

Essentially, the blurring operator K, the coefficient β_2 in $V(\boldsymbol{\beta},\boldsymbol{u})$ and the non-negativity constraint can change the global energy of the image, whereas β_1 and β_3 in $V(\boldsymbol{\beta},u)$, interpreted as scale parameters, can influence the local diffusion. See Appendix A for the proofs and analysis of these properties. The experimental effectiveness of these terms will be discussed in Section 3.1.

2.2. L-NNCV: search coefficients automatically

Given the form of the non-negativity constrained variation (2.1), we search the optimal coefficients based on supervised-learning. Namely, the coefficients are learned by some prepared samples. We prepare some pairs of training samples $(f_m,f_m^*)_{m=1,2,\cdots,M}$ where f_m is the degraded input image and f_m^* is the expected output image (ground truth), and the output of NNCV model should approximate to the ground truth. Hence, the updated image and coefficients should minimize the following objective functional

$$J(\boldsymbol{u},\boldsymbol{\beta})=\frac{1}{2}\sum_{m=1}^{M}\|u_m-f_m^*\|^2+\frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2, \tag{2.3}$$

where α_i $(i=1,2,3)$ are positive weighting coefficients related to the number of training samples and u_m $(m=1,\cdots,M)$ are the updated images computed from NNCV model. The updated output images play an important role in objective functional and inherit most features from the ground truth through the first term of objective functional (2.3), and the

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

second term ensures the convexity of objective functional [5]. The tendency of $u_m = f_m^*$ makes NNCV close to ground truth, but $\boldsymbol{\beta} = (0,0,0)$ rarely appears because u_m which minimizes NNCV model is also related to $\boldsymbol{\beta}$. To further explain how the objective functional works, we detail the learning process in iteration as follows: (1) images, updated by NNCV model, are delivered to the objective functional (2.3) to be u_m $(m = 1,\cdots,M)$; (2) the coefficients, minimizing the objective functional (2.3), are returned to NNCV model to be the coefficients in the next iteration. Then the coefficients are searched by some prepared training samples, which make NNCV image-driven. We denote this learning-based NNCV by L-NNCV as follows

$$\min_{u_m,\boldsymbol{\beta}} \quad \frac{1}{2}\sum_{m=1}^{M}\|u_m - f_m^*\|^2 + \frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2, \tag{2.4a}$$

$$\min_{u_m,u_m\geq 0} \; P(f_m,u_m) + V(\boldsymbol{\beta},u_m) \quad m = 1,\cdots,M. \tag{2.4b}$$

In this paper, we only focus on the image denoising and deblurring problems, so $P(f_m,u_m) = \frac{1}{2}\|Ku_m - f_m\|^2$.

2.3. Solving L-NNCV via ε-constraint method

The L-NNCV approach for image denoising and deblurring can be rewritten as

$$\min_{u_m,\boldsymbol{\beta}} \frac{1}{2}\sum_{m=1}^{M}\|u_m - f_m^*\|^2 + \frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2, \tag{2.5a}$$

$$\min_{u_m} \frac{1}{2}\|Ku_m - f_m\|^2 + V(\boldsymbol{\beta},u_m) \quad m = 1,\cdots,M, \tag{2.5b}$$

$$\text{s.t.} \;\; u_m \geq 0, \;\; m = 1,\cdots,M, \tag{2.5c}$$

which is a multiobjective optimization problem with non-negati-vity constraint. Non-negativity constraint can improve the quality of the solutions but complicates the variational models, so we use the classical ε-constraint technique, proposed by Haimes et al. [23], for generating Pareto optimal solution.

Then, we sketch the existing theory of ε-constraint method that we will borrow. Find an efficient point for the following minimization problem

$$\min_{\boldsymbol{\alpha},\boldsymbol{u}}[\mathscr{J}(\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{u}),\ \mathscr{H}(\boldsymbol{u})], \quad \text{s.t.}\mathscr{G}(\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{u}) \leq 0. \tag{2.6}$$

Consider

$$\min_{\boldsymbol{\alpha},\boldsymbol{u}}\mathscr{J}(\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{u}), \quad \text{s.t.}\begin{cases}\mathscr{H}(\boldsymbol{u}) \leq \varepsilon,\\ \mathscr{G}(\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{u}) \leq 0.\end{cases} \tag{2.7}$$

For convenience, we denote these problems (2.6)-(2.7) by Problem A and Problem B(ε) and let $\boldsymbol{V} = (\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{u})$.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

Theorem 2.1 (Equivalence Theorem [23]). *Let $\varepsilon \geq \min_{\boldsymbol{u}} \mathscr{H}(\boldsymbol{u})$, let $\boldsymbol{V}^*$ solve Problem $B(\varepsilon)$, and assume that, if $\boldsymbol{V}^*$ is not unique, then $\boldsymbol{V}^*$ is an optimal solution of Problem $B(\varepsilon)$ with minimal $\mathscr{H}(\boldsymbol{u})$ value. Then $\boldsymbol{V}^*$ solves Problem A.*

Obviously, ε-constraint method transforms the multiobjective problem into single-objective problem with constraints. Based on the theory, the equivalent ε-constraint problem of L-N NCV is

$$\min_{u_m,\boldsymbol{\beta}} \frac{1}{2}\sum_{m=1}^{M}\|u_m - f_m^*\|^2 + \frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2, \qquad \text{s.t.} \begin{cases} H_m(u_m,\boldsymbol{\beta}) \leq \varepsilon_m, \\ u_m \geq 0,\ m = 1, \cdots, M, \end{cases} \tag{2.8}$$

where $H_m(u_m,\boldsymbol{\beta}) := \frac{1}{2}\|Ku_m - f_m\|^2 + V(\boldsymbol{\beta}, u_m)$. The upper bound ε_m should vary to obtain subset of the Pareto optimal set. For the sake of simplicity and time saving, we force ε_m to reach almost the minimum of $H_m(u_m,\boldsymbol{\beta})$, so we plug $\nabla_{u_m} H_m(u_m,\boldsymbol{\beta}) = 0$ into (2.8). The term $\frac{\nabla u_m}{|\nabla u_m|}$ in $\nabla_{u_m} H_m(u_m,\boldsymbol{\beta})$ can be replaced by the dual variable $\boldsymbol{p}_m$ through the Fenchel dual transform, so we add $\nabla u - |\nabla u|_\tau \boldsymbol{p} = 0$ and the dual variable's constraint to (2.8). Now, we arrive at the problem

$$\min_{u_m,\boldsymbol{\beta}} \frac{1}{2}\sum_{m=1}^{M}\|u_m - f_m^*\|^2 + \frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2, \qquad \text{s.t.} \begin{cases} \nabla_{u_m} H_m(u_m,\boldsymbol{\beta}) = 0, \\ \nabla u_m - |\nabla u_m|_\tau \boldsymbol{p}_m = 0, \\ u_m \geq 0,\ \ |\boldsymbol{p}_m|^2 - 1 \leq 0, \\ m = 1, \cdots, M, \end{cases} \tag{2.9}$$

where $\nabla_{u_m} H_m(u_m,\boldsymbol{\beta}) := Au_m - K^T f_m - \beta_1 \mathrm{div}\boldsymbol{p}_m$. In order to solve the above constrained problem (2.9), one operates on the following Lagrangian functional

$$\begin{aligned} F(\boldsymbol{u},\boldsymbol{\beta}) =& \frac{1}{2}\sum_{m=1}^{M}\|u_m - f_m^*\|^2 + \frac{1}{2}\sum_{i=1}^{3}\alpha_i\beta_i^2 + \sum_{m=1}^{M}\Big[\big\langle \phi_{1m}, \nabla_{u_m} H_m\big\rangle \\ &+ \langle \phi_{2m}, \nabla u_m - |\nabla u_m|_\tau \boldsymbol{p}_m\rangle - \langle u_m,\ \mu_{1m}\rangle + \big\langle |\boldsymbol{p}_m|^2 - 1,\ \mu_{2m}\big\rangle\Big], \end{aligned} \tag{2.10}$$

where $\phi_{\cdot m}$ and $\mu_{\cdot m}$ are the multipliers and $\langle\cdot\,,\,\cdot\rangle$ is the inner product. The corresponding optimality conditions (KKT conditions [5]) for problem (2.10) are given by

$$u_m - f_m^* + \nabla_{u_m}\Big(\big\langle \phi_{1m}, \nabla_{u_m} H_m\big\rangle\Big) - \nabla\phi_{2m} + \nabla\big(\phi_{2m}\boldsymbol{p}_m^2\big) - \mu_{1m} = 0, \tag{2.11a}$$

$$Au_m - K^T f_m - \beta_1 \mathrm{div}\boldsymbol{p}_m = 0, \tag{2.11b}$$

$$\nabla u_m - |\nabla u_m|_\tau \boldsymbol{p}_m = 0, \tag{2.11c}$$

$$\big(|\boldsymbol{p}_m|^2 - 1\big) \odot \mu_{2m} = 0, \tag{2.11d}$$

$$u_m \odot \mu_{1m} = 0, \tag{2.11e}$$

$$\mu_{1m},\ \ \mu_{2m} \geq 0. \tag{2.11f}$$

Here, $\odot$ is understood as component-wise multiplication. As for $\beta_i (i = 1, 2, 3)$, we compute the derivative of $F(\boldsymbol{u},\boldsymbol{\beta})$ with respect to them. Instead of solving the optimality conditions (2.11) derived from (2.9) directly, we simplify (2.9) further to reduce complexity. After

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

being computed by NNCV, u_m $(m = 1, \cdots, M)$ satisfies the non-negativity constraint and the Fenchel dual transform, so these constraints will vanish in (2.9) and $\nabla_{u_m} H_m(u_m, \boldsymbol{\beta})$ should be replaced by $\nabla_{u_m} L_m := Au_m - K^T f_m - \beta_1 \mathrm{div}(\boldsymbol{p}_m) - \lambda_m = 0$where λ_m is the multiplier of non-negativity constraint. Therefore, we solve the following minimization problem

$$\min_{u_m, \boldsymbol{\beta}} \frac{1}{2} \sum_{m=1}^{M} \|u_m - f_m^*\|^2 + \frac{1}{2} \sum_{i=1}^{3} \alpha_i \beta_i^2, \quad \text{s.t. } \nabla_{u_m} L_m(u_m, \boldsymbol{\beta}) = 0,\ m = 1, \cdots, M. \tag{2.12}$$

Then the corresponding Lagrangian functional is given by

$$\tilde{F}(\boldsymbol{u}, \boldsymbol{\beta}) = \frac{1}{2} \sum_{m=1}^{M} \|u_m - f_m^*\|^2 + \frac{1}{2} \sum_{i=1}^{3} \alpha_i \beta_i^2 + \sum_{m=1}^{M} \left\langle \phi_m, \nabla_{u_m} L_m \right\rangle,$$

where ϕ_m is the corresponding Lagrange multiplier. To minimize $\tilde{F}$, the derivative $\frac{\partial \tilde{F}}{\partial u_m}$ is computed by

$$\frac{\partial \tilde{F}}{\partial u_m} = u_m - f_m^* + (K^T K + \beta_2 I)\phi_m - \beta_3 \Delta \phi_m. \tag{2.13}$$

Additionally,

$$\frac{\partial \tilde{F}}{\partial \beta_1} = \alpha_1 \beta_1 - \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot \mathrm{div} \boldsymbol{p}_m dx, \tag{2.14a}$$

$$\frac{\partial \tilde{F}}{\partial \beta_2} = \alpha_2 \beta_2 + \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot u_m dx, \tag{2.14b}$$

$$\frac{\partial \tilde{F}}{\partial \beta_3} = \alpha_3 \beta_3 - \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot \Delta u_m dx. \tag{2.14c}$$

In conclusion, the updates of L-NNCV are as follows: first, the image u_m is updated by NNCV model with coefficients of the former iteration; then, the coefficients are updated by solving

$$u_m - f_m^* + (K^T K + \beta_2 I)\phi_m - \beta_3 \Delta \phi_m = 0, \tag{2.15a}$$

$$\alpha_1 \beta_1 - \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot \mathrm{div} \boldsymbol{p}_m dx = 0, \tag{2.15b}$$

$$\alpha_2 \beta_2 + \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot u_m dx = 0, \tag{2.15c}$$

$$\alpha_3 \beta_3 - \sum_{m=1}^{M} \int_{\Omega} \phi_m \odot \Delta u_m dx = 0. \tag{2.15d}$$

Through the learning process, NNCV becomes $J(\boldsymbol{u}, \boldsymbol{\beta})$-driven which can be considered as image-driven.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

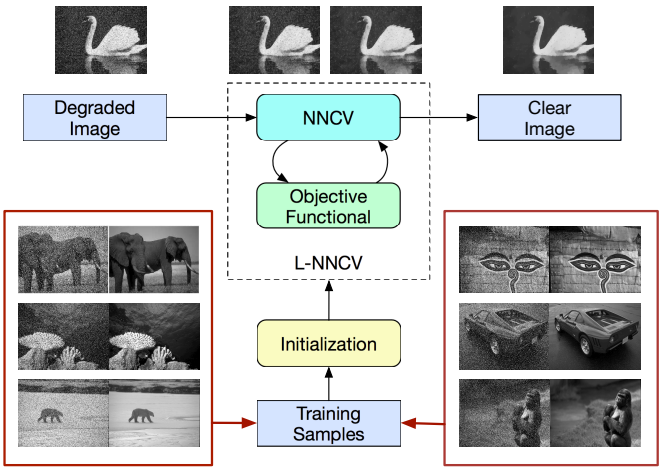


Figure 1: The proposed framework including NNCV model and L-NNCV approach.

Algorithm 2.1 Non-Negativity Constrained Variation (NNCV)

Input: Blurring image f.

Output: Recovered image u.

1: Initialize $u^0, \boldsymbol{p}^0, \lambda^0$. Set $k = 0$.
2: **while** not converge **do**
3: Find the inactive set $\mathscr{I}^k$ and the active set $\mathscr{A}^k$.
4: Compute $\delta\lambda^k_{\mathscr{A}}$ (the search direction of λ_m on the active set).
5: Compute $\delta\boldsymbol{p}^k$ (the search direction of $\boldsymbol{p}_m$).
6: Compute $\delta u^k_{\mathscr{A}}$ (the search direction of u_m on the active set).
7: Update the components in $\mathscr{I}^k$ and $\mathscr{A}^k$ by (B.19).
8: **end while**

2.4. The framework for image denoising and deblurring

The framework for image denoising and deblurring is summarized in Fig. 1. At first, some pairs of training samples, including degraded input images and the desired output images, are prepared as ones in red circles. Then the coefficients are initialized by solving (2.15a)-(2.15d). The initialized coefficients are delivered to L-NNCV to search the optimal coefficients.

There are several methods to handle NNCV model such as proximal splitting [2,4] and primal-dual active-set strategy [9,24,25]. For high convergence rate, we use primal-dual active-set strategy to solve NNCV model. See Appendix B for all derivatives of primal-dual active-set strategy for NNCV model. The algorithms of NNCV and L-NNCV are shown in Algorithm 2.1 and 2.2, respectively.

Last but not the least, the framework can be applied to other image processing problems by associating variational method and objective functional.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

Algorithm 2.2 Learning-based Non-Negativity Constrained Variation (L-NNCV)

Input: Initialized coefficients $\boldsymbol{\beta}^0$ and some pairs of training image samples (f_m, f_m^*).
Output: The coefficients $\boldsymbol{\beta}$.

Initialize $u_m^0, \boldsymbol{p}_m^0, \lambda_m^0$. Set $k = 0$.
while not converged **do**
  Set $m = 1$.
  while $m \leq M$ **do**
    Find the inactive set $\mathscr{I}_m^k$ and the active set $\mathscr{A}_m^k$.
    Compute $\delta\lambda_{m.\mathscr{A}_m}^k$ (the search direction of λ_m on the active set).
    Compute $\delta\boldsymbol{p}_m^k$ (the search direction of $\boldsymbol{p}_m$).
    Compute $\delta u_{m.\mathscr{A}_m}^k$ (the search direction of u_m on the active set).
    Update the components in $\mathscr{I}_m^k$ and $\mathscr{A}_m^k$ by (B.19).
    Compute ϕ_m, using (2.15a).
  end while
  Update the coefficients in the next iteration $\boldsymbol{\beta}^{k+1}$, using (2.15b)-(2.15d).
end while

2.5. Differences between traditional TV models, L-PDE and our proposed model

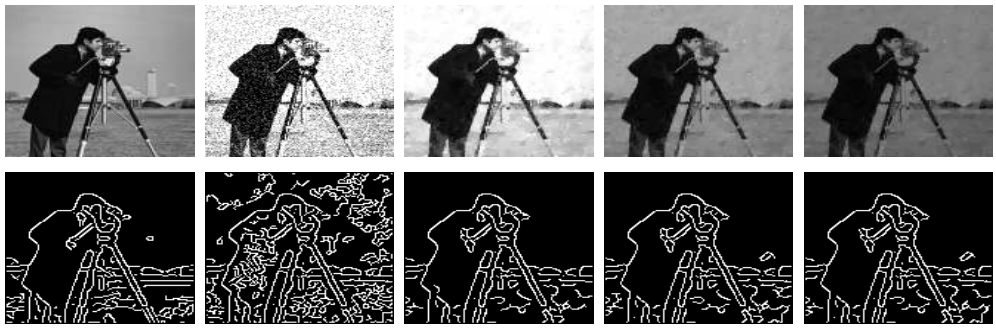
Intrinsically, NNCV model is automatically driven by the real image information, however, traditional TV models are manually designed by the desired properties. The learned coefficients dynamically vary to adapt special images under different conditions. Clearly, some of traditional TV models, such as TV-based deblurring model, are special cases of our proposed model. It is reasonable to apply learning-based idea to other traditional TV models. In addition, the learning-based idea makes the variation more flexible to different issues with the help of the prepared training images.

In contrast to L-PDE, NNCV model considers the non-negativity constraint. As stated in Section 1, non-negativity constraint not only is the requirement of some problems but also can achieve better performance, but it is difficult to handle, so we borrow the existing ε-constraint technique to solve it. Moreover, NNCV model and L-PDE stand at different points of the problem. Basically, L-PDE considers the diffusion process which needs proficient understanding of operators, while NNCV model contains energy functional which can collect our desired properties together.

3. Experimental results

In this section, we show some experimental results in three aspects: effectiveness of each term in the image-driven regularization, comparisons on image denoising and deblurring. The peak signal-to-noise ratio (PSNR), defined by $10\log_{10}\left(\frac{255^2}{\frac{1}{mn}\|original-reconstructed\|^2}\right)$, measures image reconstruction error.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press



(a) - (b) 10.61dB (c) 10.79dB (d) 27.14dB (e) 18.12dB

Figure 2: The results of denoising images with brightness and Gaussian noise. (a) Original image. (b) Noisy exposed image. (c)-(e) Recovered images for β_2 of 0, 0.55 and 1, respectively. The corresponding edges are detected by Canny algorithm [7] with threshold of $[0, 0.1]$.

Table 1: Performance comparison of PN [44], CGM [12], NNCGM [25] and NNCV on denoising images with Gaussian and Salt & Pepper noise when the optimal coefficients are chosen. ND means density of the Salt & Pepper noise.

SNR	ND	PN	CGM	NNCGM	NNCV
15	0.005	20.34	26.19	26.19	**26.26**
	0.01	19.63	25.06	25.09	**25.46**
19	0.005	23.00	26.66	26.66	**27.12**
	0.01	22.03	25.58	25.58	**26.34**
23	0.005	25.71	27.99	27.99	**28.56**
	0.01	23.58	25.76	25.76	**26.93**

3.1. Effectiveness of each term in image-driven regularization

We analyze the effectiveness of each term in the image-driven regularization in NNCV model (2.2) experimentally.

Fig. 2 demonstrates the results of NNCV model for a fixed β_1 and different β_2. The noisy image is deteriorated by brightness and Gaussian noise with a fixed SNR of 23 dB. We can see that v_2 in the image-driven regularization darkens the image, which means that the coefficient β_2 can affect the global energy of the image as stated in Section 2.1. However, the edges of denoised images detected by Canny algorithm [7] are close to the original image. Namely, v_2 term can hardly influence the local diffusion while it penalizes brightness of images without changing the local information such as gradient.

As is known to all, v_1 term (also known as total variation) keeps good balance of smoothness, so it can not only remove noise (especially Gaussian noise) but also preserve edges [1]. v_3 term in the image-driven regularization usually smoothes the image too much due to the mightiness of l_2 norm. However, it will show better performance when the image is deteriorated by some strong noise such as salt & pepper noise. As a result, we examine the performance of v_1 and v_3 in the image-driven regularization on suppressing these noise. The PSNRs of PN [44], CGM [12], NNCGM [25] and proposed NNCV are

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press



(a) – (b) 20.14dB (c) 21.02dB (d) 25.84dB (e) 25.97dB (f) 27.56dB



(g) 20.69dB (h) 11.58dB (i) 21.05dB (j) 26.68dB (k) 26.68dB (l) 27.58dB

Figure 3: The results of denoising images with Gaussian noise. (a) Original image. (g) Noisy image with a fix SNR of 15dB. (b)-(f) (h)-(l) Denoised images using AM [25], PN [44], CGM [12], NNCGM [25] and L-NNCV by column, respectively. The coefficients of TV norm are 9 (top) and 25 (bottom).

shown in Table 1 when Gaussian and salt & pepper noise are added to the image. Here, ND means density of the salt & pepper noise. Obviously, NNCV reaches highest PSNRs.

In conclusion, the experimental effectiveness of each term in the image-driven regularization accords with the theoretical properties in Section 2.1. Each term adapts different images under different conditions, which makes the adaption of NNCV model possible.

3.2. Comparisons on image denoising

The qualitative performance of AM [25], PN [44], CGM [12], NNCGM [25] and L-NNCV is presented in Fig. 3 for different regulated coefficients (9 or 25) of TV norm with a fixed SNR of 15dB. Comparatively speaking, L-NNCV can reach highest PSNRs no matter how the coefficient of TV is initialized. Substantially, the assumption of image-driven and adaptive variation keeps good balance between under-regularization and over-regularization. The AM and PN methods are available to solve the non-negativity constrained TV-based problems theoretically, whereas they hardly produce a good result in a limited iterations due to its slow speed of convergence in practice. The CGM, NNCGM and L-NNCV usually converge in 20 iterations and have better performance.

Table 2 shows the quantitative results of AM, PN, CGM, NNCGM and L-NNCV when the optimal coefficients are chosen for varying SNRs and test images. The subrows in each row are Cameraman (C) and Satellite (S) data. Obviously, L-NNCV slightly surpasses NNCGM which is superior to other methods.

To further testify to the workable L-NNCV, we perform experiments on four databases: BSDS500 [30], LIVE [39], CSIQ [27] and TID2013 [35] including training samples and test images. On BSDS500 [30], the images are randomly separated into 6 teams and we perform two experiments with synthetic noise including zero-mean white Gaussian noise ($\sigma = 25$) and a mixture of zero-mean Gaussian white noise ($\sigma = 50$), Poisson noise and salt & pepper noise ($d = 0.1$). On LIVE [39], CSIQ [27] and TID2013 [35], we perform experiments with various types of noise with 5 levels including white Gaussian

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

Table 2: Performance comparison of AM [25], PN [44], CGM [12], NNCGM [25] and L-NNCV on denoising images with Gaussian noise when the optimal coefficients are chosen. The subrows in each row are Cameraman (C) and Satellite (S) data.

SNR	Im	AM	PN	CGM	NNCGM	L-NNCV
15	C	21.98	21.02	27.57	27.58	**27.60**
	S	29.95	32.01	36.61	37.07	**37.16**
19	C	26.29	21.02	29.23	30.12	**30.16**
	S	34.57	36.00	39.82	40.34	**40.43**
23	C	30.70	28.80	29.6	**32.90**	**32.90**
	S	38.99	39.98	43.02	43.83	**43.93**

Table 3: Average PSNRs of noisy images and denoised images by L-PDE [29], L-NNCV, NNCGM [25] and NNCV-F on BSDS500 [30], LIVE [39], CSIQ [27] and TID2013 [35].

Database	Noise	Average PSNRs				
		Noisy Images	L-PDE	L-NNCV	NNCGM	NNCV-F
LIVE	WGN	23.51	-	28.12	28.39	**28.47**
CSIQ	WGN	31.63	34.12	32.23	**34.71**	**34.71**
BSDS500	GAU	20.46	25.60	27.31	27.48	**27.49**
	MIX	12.12	21.63	21.61	21.82	**21.90**
TID2013	WGN	30.63	33.43	33.96	**34.07**	**34.07**
	HFN	28.23	32.57	32.94	**33.02**	**33.02**
	MGN	30.15	32.62	32.97	**33.11**	**33.11**
	MN	30.33	31.85	32.32	32.49	**32.60**
	SCN	30.63	31.43	31.93	**32.12**	**32.12**

noise (WGN), high frequency noise (HFN), multiplicative Gaussian noise (MGN), masked noise (MN) and spatially correlated noise (SCN).

Table 3 shows the average PSNRs of noisy images and denoised images by L-PDE [29], L-NNCV, NNCGM [25], NNCV-F[†] and Figs. 4-5 give qualitative comparisons. Obviously, NNCV-F achieves higher average PSNRs. Especially, NNCV-F have better performance when dealing with strong noise due to the addition of l_2 norm.

Moreover, we analyze the relation between the searched coefficients of L-NNCV and the optimal coefficients of NNCV-F. Fig. 6 presents the coefficients with varying training samples and initialized coefficients. It is obvious that L-NNCV can search for the optimal coefficient of NNCV-F automatically, so we denote the searched coefficients of L-NNCV by quasi-optimal coefficients. Besides, we can see that L-NNCV is independent with the choice of training samples.

[†]The coefficients of NNCV is initialized optimally and fixed in each iteration. At this situation, NNCV is denoted as NNCV-F briefly.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

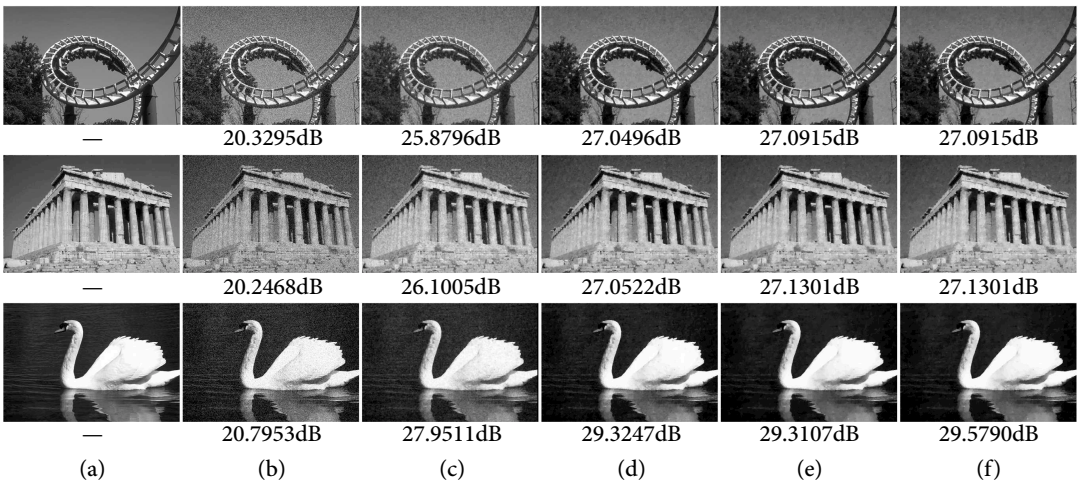


Figure 4: The results of denoising images with Gaussian noise on database from BSDS500 [30]. (a) Original images. (b) Noisy images with Gaussian noise. (c)-(f) Denoised images using L-PDE [29], L-NNCV, NNCGM [25] and NNCV-F, respectively.

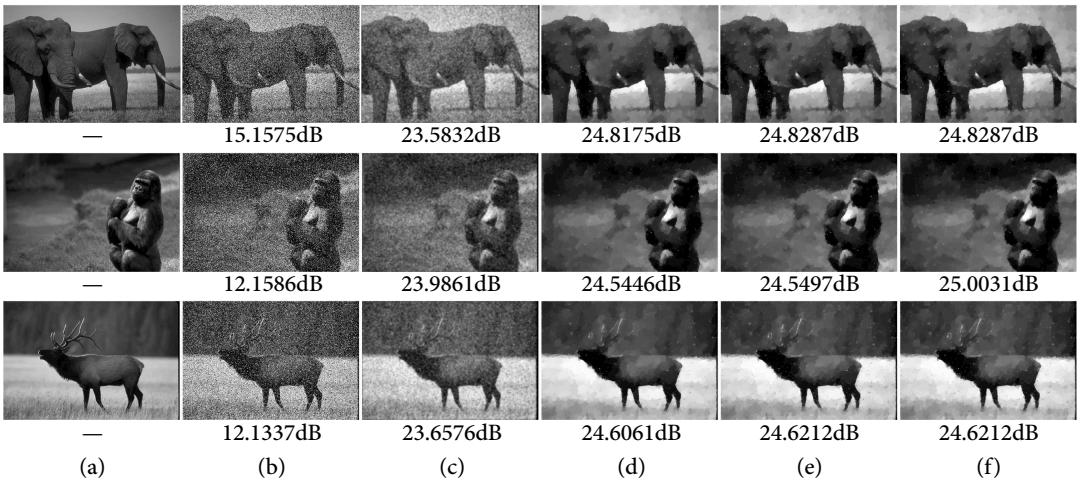


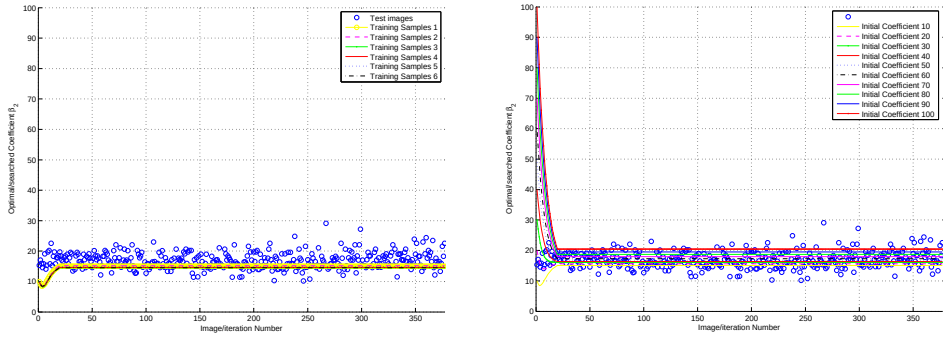
Figure 5: The results of denoising images with mixture noise on database from BSDS500 [30]. (a) Original images. (b) Noisy images with mixture noise. (c)-(f) Denoised images using L-PDE [29], L-NNCV, NNCGM [25] and NNCV-F, respectively.

3.3. Comparisons on image deblurring

For image deblurring, we compare L-NNCV with AM [25], PN [44], CGM [12] and NNCGM [25], on images with varying SNRs and sizes of PSF. Table 4 gives the quantitative results on Cameraman image. It can be seen that L-NNCV can overcome NNCGM which outperforms other methods. Besides, L-NNCV can search for the quasi-optimal coefficients automatically in contrast with NNCGM.

The robustness of the L-NNCV is tested with respect to the smoothing parameter τ in

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press



(a) (b)

Figure 6: Optimal coefficients of NNCV-F and searched coefficients of L-NNCV with varying training samples (a) and initialization (b).

Table 4: Performance comparison of AM [25], PN [44], CGM [12], NNCGM [25] and L-NNCV on deblurring when the optimal coefficients are chosen.

PSF	SNR	AM	PN	CGM	NNCGM	L-NNCV
3×3	15	–	19.76	24.80	24.85	**24.89**
	19	24.30	22.44	25.95	25.98	**26.03**
	23	26.96	24.57	27.05	27.10	**27.13**
9×9	15	20.28	18.23	21.59	21.60	**21.70**
	19	22.11	20.05	22.35	22.36	**22.40**
	23	22.70	21.12	23.04	23.04	**23.09**

(B.12a) and the constant c in (B.12c). Figs. 7-8 present PSNRs and KKT residuals of L-NNCV for varying c and τ with a fixed SNR of 17dB and Cameraman image. Obviously, c and τ do not influence PSNRs in every iteration. The constant c does not influence the KKT residual after 15 iterations, and KKT residual is robust with the parameter τ below 0.001.

4. Discussion

In this paper, we propose a heuristic learning-based non-negativity constrained variation including a problem-based term and an image-driven regularization for image denoising and deblurring. The image-driven regularization is learned by prepared training samples to adapt different images. The proposed learning-based variation is solved by ε-constraint method when applied on image denoising and deblurring. The non-negativity constrained variation model is solved by primal-dual active-set method where the techniques inherit from Krishnan et al. [25]. The learning-based approach can be more educated than wild guess in choosing the coefficients for regularization terms. It works more efficiently on images with similar degraded pattern. The choice of training samples influences the effectiveness of this approach. Experimental results show the effectiveness of our

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

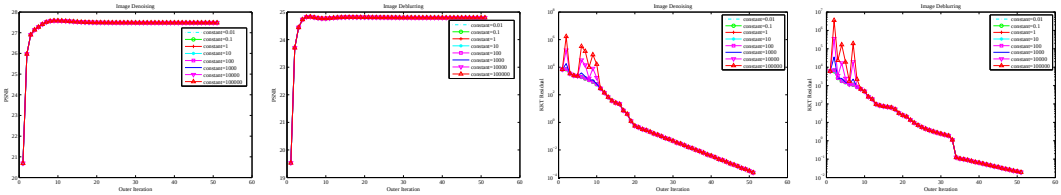


Figure 7: PSNRs (top) and KKT (bottom) residuals of L-NNCV for varying c on image denoising (left) and deblurring (right).

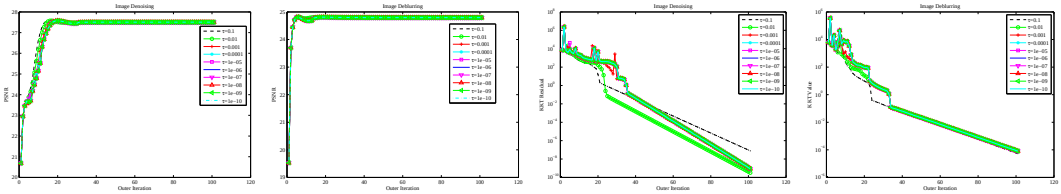


Figure 8: PSNRs (top) and KKT residuals (bottom) of L-NNCV for varying τ on image denoising (left) and deblurring (right).

methods compared with traditional TV-based models and L-PDE.

The future work will focus on the following aspects. First, L-NNCV can search for global quasi-optimal coefficients, but the coefficients are not optimal for every image which sometimes results in weak trade-off in different image locations, so we look forward to eliminating the loss of effectiveness. Second, we would like to enrich the image-driven regularization to make it more adaptive to different problems. Finally, we will extend the applications of our framework to image segmentation, inpainting and other image processing problems.

Appendices

A. Interpretation and properties of coefficients

In this section, we give some elementary interpretation of each term in the image-driven regularization and the scale properties of the coefficients. Assuming that $u(\cdot, \boldsymbol{\beta})$ is the minimizer of NNCV denoising and deblurring model (B.8) and $u \in \mathbb{W}^{1,2}(\Omega)$ where $\mathbb{W}^{1,2}(\Omega) = \{v : v \in \mathbb{L}^2(\Omega), \nabla v \in \mathbb{L}^2(\Omega)\}$, we give Theorem A.1 and A.2 for interpretation of the coefficients.

Theorem A.1. *$\|Ku\|^2 \leq C$ where C is independent of $\boldsymbol{\beta}$.*

Proof. For arbitrary $v \in \mathbb{W}^{1,2}(\Omega)$, we have

$$\frac{1}{2}\|Ku - f\|^2 + V(u, \boldsymbol{\beta}) \leq \frac{1}{2}\|Kv - f\|^2 + V(v, \boldsymbol{\beta}). \tag{A.1}$$

By setting $v = 0$, then

$$\|Ku - f\|^2 \leq \|f\|^2,$$

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

which results in

$$\|Ku\|^2 \leq 2\|f\|^2.$$

We denote $C = 2\|f\|^2$ so as to complete the proof. □

Theorem A.2. *If $K = I$, $\beta_2 = 0$ and the non-negativity constraint is left out, u converges to a constant:* $\lim_{\beta_1 \to \infty} \|u - \bar{f}\| = 0$, $\lim_{\beta_3 \to \infty} \|u - \bar{f}\|^2 = 0$ *where* $\bar{f} = \frac{1}{|\Omega|}\int_\Omega f(\boldsymbol{x})d\boldsymbol{x}$.

Proof. By setting $v = 0$ in (A.1), we obtain

$$0 \leq \beta_1 \|u\|_{TV} \leq \frac{1}{2}\|Ku - f\|^2 + V(u, \boldsymbol{\beta}) \leq \frac{1}{2}\|f\|^2 \tag{A.2}$$

from which we can deduce $\lim_{\beta_1 \to \infty} \|\nabla u\| = 0$. When applying primal-dual active-set strategy to NNCV, the following equation holds

$$\left(K^T K + \beta_2 I\right)u = K^T f + \operatorname{div}(B\nabla u) + \lambda, \tag{A.3}$$

where $B := \beta_1 \frac{1}{|\nabla u|} + \beta_3$. By integrating on the region Ω, the divergence operator will vanish by the Green formula as follows

$$\begin{aligned}\int_\Omega \left(K^T K + \beta_2 I\right)u d\boldsymbol{x} &= \int_\Omega \operatorname{div}(B\nabla u)d\boldsymbol{x} + \int_\Omega K^T f\, d\boldsymbol{x} + \lambda|\Omega| \\ &= \int_{\partial\Omega} B\frac{\partial u}{\partial N}d\boldsymbol{x} + \int_\Omega K^T f\, d\boldsymbol{x} + \lambda|\Omega| \\ &= \int_\Omega K^T f\, d\boldsymbol{x} + \lambda|\Omega|. \end{aligned} \tag{A.4}$$

Moreover,

$$\|u - \bar{f}\| = \left\|u - \frac{1}{|\Omega|}\int_\Omega f(\boldsymbol{x})d\boldsymbol{x}\right\|. \tag{A.5}$$

When $K = I$, we can plug (A.4) into (A.5) to obtain

$$\|u - \bar{f}\| = \int_\Omega \left|u - (\beta_2 + 1)\bar{u} + \lambda\right| d\boldsymbol{x}. \tag{A.6}$$

If $\beta_2 = 0$ and the non-negativity constraint is left out, we can use the Poincaré inequality to transform (A.6) into

$$\|u - \bar{f}\| \leq C\|\nabla u\|. \tag{A.7}$$

Therefore, $\lim_{\beta_1 \to \infty} \|u - \bar{f}\| = 0$. Similarly, we can deduce that $\lim_{\beta_3 \to \infty} \|u - \bar{f}\|^2 = 0$. □

Meanwhile, equation $\int_\Omega u(\boldsymbol{x})d\boldsymbol{x} = \int_\Omega f(\boldsymbol{x})d\boldsymbol{x}$ can be derived from (A.4) if the conditions of Theorem A.2 are satisfied. Moreover, we can deduce $(1 + \beta_2)\int_\Omega u(\boldsymbol{x})d\boldsymbol{x} = \int_\Omega f(\boldsymbol{x})d\boldsymbol{x}$ from (A.4). From these properties and proofs, we can see that the blurring operator, β_2 and the non-negativity constraint influence the global energy of an image, whereas β_1 and β_3 influence the local diffusion. Coefficients β_1 and β_3 can be understood as scale parameters which control the constructed family of images $u(\boldsymbol{x}, \beta_i)$ $(i = 1, 3)$ started from the blurring image f [1, 42, 45].

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

B. Primal-dual active-set strategy for NNCV model

The non-negativity constrained variation (NNCV) denoising and deblurring model is

$$\min_{u,u\geq 0} \frac{1}{2}\|Ku-f\|^2 + V(\boldsymbol{\beta},u), \tag{B.8}$$

where $V(\boldsymbol{\beta},u) := \beta_1 v_i(u) + \frac{1}{2}\beta_2 v_2(u) + \frac{1}{2}\beta_3 v_3(u)$. It can be rewritten as

$$\min_{u} \frac{1}{2}\|Ku-f\|^2 + V(\boldsymbol{\beta},u) + I(u\geq 0), \tag{B.9}$$

where $I(u \geq 0) = 0$ if $u \geq 0$ for all components of u and $I(u \geq 0) = \infty$ if any of the components of u is < 0. The Lagrangian functional for problem (B.9) is given by

$$L(u,\lambda) := \|Ku-f\|^2 + V(\boldsymbol{\beta},u) - \langle u,\lambda\rangle,$$

where λ is the Lagrange multiplier which handles $u \geq 0$. Then we have

$$\nabla_u L = Au - K^T f - \beta_1 \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \lambda, \tag{B.10}$$

where $A := K^T K + \beta_2 I - \beta_3 \Delta$. Owing to the presence of the term $\frac{1}{|\nabla u|}$, $|\nabla u|$ should be slightly perturbed to make (B.10) well-defined. Another difficulty caused by (B.10) is the linearization of the highly nonlinear term $\mathrm{div}(\frac{\nabla u}{|\nabla u|})$, which can be overcome with the help of an auxiliary variable $\boldsymbol{p} = \frac{\nabla u}{|\nabla u|}$ proposed by Chan et al. [12]. So the Karush-Kuhn-Tucker (KKT) conditions [5] for problem (B.8) are as follows

$$Au - K^T f - \beta_1 \mathrm{div}\boldsymbol{p} - \lambda = 0, \tag{B.11a}$$
$$\nabla u - |\nabla u|_\tau \boldsymbol{p} = 0, \tag{B.11b}$$
$$u \odot \lambda = 0, \tag{B.11c}$$
$$\lambda,\ u \geq 0. \tag{B.11d}$$

Here, equation $u \odot \lambda = 0$ is understood as component-wise multiplication. In order to express (B.11c) and (B.11d) as a single equality constraint, we replace them by $\lambda - \max\{0,\ \lambda - cu\} = 0$ where c is an arbitrary positive constant. Then we have the following equivalent system of equations:

$$\nabla u - |\nabla u|_\tau \boldsymbol{p} = 0, \tag{B.12a}$$
$$Au - K^T f - \beta_1 \mathrm{div}\boldsymbol{p} - \lambda = 0, \tag{B.12b}$$
$$\lambda - \max\{0,\ \lambda - cu\} = 0. \tag{B.12c}$$

For convenience, the left hand sides of (B.12a)-(B.12c) are denoted by $F_1(\boldsymbol{p},u,\lambda)$, $F_2(\boldsymbol{p},u,\lambda)$ and $F_3(\boldsymbol{p},u,\lambda)$ respectively. The semi-smooth Newton's update for the system (B.12a)-(B.12c) is given by

$$\begin{bmatrix} |\nabla u|_\tau & -\left(I - \frac{\boldsymbol{p}(\nabla u)^T}{|\nabla u|_\tau}\right)\nabla & 0 \\ -\beta_1 \mathrm{div} & A & -I \\ 0 & \frac{\partial F_3}{\partial u} & \frac{\partial F_3}{\partial \lambda} \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{p} \\ \delta u \\ \delta \lambda \end{bmatrix} = - \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}. \tag{B.13}$$

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

Instead of solving the above system directly, we use the active-set idea [24] to handle the non-negativity constraint. After the k-th Newton step of (B.13), the predicted inactive and active sets are defined by $\mathscr{I}^k := \{i : \lambda_i^k - cu_i^k \le 0\}$ and $\mathscr{A}^k := \{i : \lambda_i^k - cu_i^k > 0\}$ respectively. The components of u in $\mathscr{I}$ and $\mathscr{A}$ are computed by the corresponding down-sampling matrix $D_{\mathscr{I}}$ and $D_{\mathscr{A}}$, for example, $u_{\mathscr{I}} = D_{\mathscr{I}}u$. Using the active and inactive set gives

$$\begin{bmatrix} |\nabla u|_\tau & Q\nabla D_{\mathscr{I}}^T & Q\nabla D_{\mathscr{A}}^T & 0 & 0 \\ -\beta_1 D_{\mathscr{I}}\mathrm{div} & A_{\mathscr{I}} & A_{\mathscr{I}\mathscr{A}} & -I & 0 \\ -\beta_1 D_{\mathscr{A}}\mathrm{div} & A_{\mathscr{A}\mathscr{I}} & A_{\mathscr{A}} & 0 & -I \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & cI & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \delta \boldsymbol{p} \\ \delta u_{\mathscr{I}} \\ \delta u_{\mathscr{A}} \\ \delta \lambda_{\mathscr{I}} \\ \delta \lambda_{\mathscr{A}} \end{bmatrix} = - \begin{bmatrix} F_1 \\ D_{\mathscr{I}}F_2 \\ D_{\mathscr{A}}F_2 \\ D_{\mathscr{I}}F_3 \\ D_{\mathscr{A}}F_3 \end{bmatrix}, \tag{B.14}$$

where $Q = -\big(I - \frac{\boldsymbol{p}(\nabla u)^T}{|\nabla u|_\tau}\big)$. Eq. (B.14) implies $\lambda_{\mathscr{I}}^{k+1} = 0$ and $u_{\mathscr{A}}^{k+1} = 0$ where the subscripts mean components in the corresponding set. By setting $\delta\lambda_{\mathscr{I}} = -\lambda_{\mathscr{I}}^k$ and $\delta u_{\mathscr{A}} = -u_{\mathscr{A}}^k$, the semi-smooth Newton step is reduced to

$$\begin{bmatrix} |\nabla u|_\tau & Q\nabla D_{\mathscr{I}}^T & 0 \\ -\beta_1 D_{\mathscr{I}}\mathrm{div} & A_{\mathscr{I}} & 0 \\ -\beta_1 D_{\mathscr{A}}\mathrm{div} & A_{\mathscr{A}\mathscr{I}} & -I \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{p} \\ \delta u_{\mathscr{I}} \\ \delta \lambda_{\mathscr{A}} \end{bmatrix} = - \begin{bmatrix} F_1 \\ D_{\mathscr{I}}F_2 \\ D_{\mathscr{A}}F_2 \end{bmatrix} + \begin{bmatrix} Q\nabla D_{\mathscr{A}}^T u_{\mathscr{A}} \\ A_{\mathscr{I}\mathscr{A}}u_{\mathscr{A}} - \lambda_{\mathscr{I}} \\ A_{\mathscr{A}}u_{\mathscr{A}} \end{bmatrix}, \tag{B.15}$$

which can be also expressed as

$$\delta\lambda_{\mathscr{A}} = -\beta_1 D_{\mathscr{A}}\mathrm{div}\delta\boldsymbol{p} + A_{\mathscr{A}\mathscr{I}}\delta u_{\mathscr{I}} + D_{\mathscr{A}}F_2 - A_{\mathscr{A}}u_{\mathscr{A}}, \tag{B.16a}$$

$$\delta\boldsymbol{p} = \frac{1}{|\nabla u|_\tau}\big[-Q\nabla(D_{\mathscr{I}}^T\delta u_{\mathscr{I}} - D_{\mathscr{A}}^T\delta u_{\mathscr{A}})\big] - F_1. \tag{B.16b}$$

After pluging (B.16a) and (B.16b) into the Newton's update (B.15) and symmetrizing the system, we update $\delta u_{\mathscr{I}}$ by solving

$$D_{\mathscr{I}}\Big[-\beta_1\mathrm{div}\frac{1}{|\nabla u|_\tau}\Big(I - \frac{\boldsymbol{p}(\nabla u)^T + (\nabla u)\boldsymbol{p}^T}{2|\nabla u|_\tau}\Big)\nabla + A\Big] \times D_{\mathscr{I}}^T\delta u_{\mathscr{I}} = G, \tag{B.17}$$

where

$$\begin{aligned} G := & -D_{\mathscr{I}}F_2 + A_{\mathscr{I}\mathscr{A}}u_{\mathscr{A}} - \lambda_{\mathscr{I}} \\ & - \beta_1 D_{\mathscr{I}}\mathrm{div}\frac{1}{|\nabla u|_\tau}\left[\Big(I - \frac{\boldsymbol{p}(\nabla u)^T + (\nabla u)\boldsymbol{p}^T}{2|\nabla u|_\tau}\Big)\nabla D_{\mathscr{A}}^T\delta u_{\mathscr{A}} + F_1\right]. \end{aligned} \tag{B.18}$$

Obviously, these derivations are highly similar to the solution process of non-negativity constrained TV deblurring problem [25], while the matrix A is different. With the help of primal-dual active-set strategy, the non-negativity constrained variation can be solved if the coefficients in the image-driven regularization are given. The components in $\mathscr{I}^k$ and

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

$\mathscr{A}^k$ are updated by

$$u_{\mathscr{A}}^{k+1} = 0, \tag{B.19a}$$
$$\lambda_{\mathscr{J}}^{k+1} = 0, \tag{B.19b}$$
$$\boldsymbol{p}^{k+1} = \boldsymbol{p}^k + s\delta\boldsymbol{p}^k, \tag{B.19c}$$
$$u_{\mathscr{J}}^{k+1} = u_{\mathscr{J}}^k + \delta u_{\mathscr{J}}^k, \tag{B.19d}$$
$$\lambda_{\mathscr{A}}^{k+1} = \lambda_{\mathscr{A}}^k + \delta\lambda_{\mathscr{A}}^k, \tag{B.19e}$$

where $\delta(\cdot)$ is the corresponding search direction computed by (B.16a)-(B.17). The step size s ensures the dual variable's quadratic constraint $|\boldsymbol{p}_{i,j}| \leq 1$ where $\boldsymbol{p}_{i,j}$ denotes any component of $\boldsymbol{p}$.

Acknowledgments The authors would like to thank the editor and reviewers for their insightful and constructive comments. The work of authors was partly supported by National Natural Science Foundation of China (Nos. 31302182, 11171374 and 61301240) and Natural Science Foundation of Shandong Province (No. ZR2011AZ001).

References

[1] G. Aubert and P. Kornprobst, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, Springer-Verlag, New York, USA, 2006.
[2] H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke and H. Wolkowicz, Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer Science and Business Media, 2011.
[3] A. Beck and M. Teboulle, *Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems*, IEEE Trans. Image Process., 18 (2009), pp. 2419–2434.
[4] S. Becker and J. Fadili, *A quasi-newton proximal splitting method*, in NIPS, 2012, pp. 2618–2626.
[5] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, New York, USA, 2004.
[6] J. F. Cai, R. H. Chan and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149.
[7] J. Canny, *A computational approach to edge detection*, IEEE Trans. Pattern Anal. Mach. Intell., 8 (1986), pp. 679–698.
[8] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
[9] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vis., 40 (2011), pp. 120–145.
[10] R. H. Chan, Y. W. Wen and A. M. Yip, *A fast optimization transfer algorithm for image inpainting in wavelet domains*, IEEE Trans. Image Process., 18 (2009), pp. 1467–1476.
[11] T. F. Chan, S. Esedoglu and M. Nikolova, *Algorithms for finding global minimizers of image segmentation and denoising models*, SIAM J. Appl. Math., 66 (2006), pp. 1632–1648.
[12] T. F. Chan, G. H. Golub and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20 (1999), pp. 1964–1977.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

[13] T. F. Chan, M. K. Ng, A. C. Yau and A. M. Yip, *Superresolution image reconstruction using fast inpainting algorithms*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 3–24.
[14] T. F. Chan and J. Shen, *Variational image inpainting*, Commun. Pure Appl. Math., 58 (2005), pp. 579–619.
[15] T. F. Chan, J. Shen and H. M. Zhou, *Total variation wavelet inpainting*, J. Math. Imaging Vis., 25 (2006), pp. 107–125.
[16] T. F. Chan and L. A. Vese, *Active contours without edges*, IEEE Trans. Image Process., 10 (2001), pp. 266–277.
[17] R. Chartrand and B. Wohlberg, *Total-variation regularization with bound constraints*, in IEEE ICASSP, 2010, pp. 766–769.
[18] X. Chen, M. K. Ng and C. Zhang, *Non-Lipschitz l_p-regularization and box constrained model for image restoration*, IEEE Trans. Image Process., 21 (2012), pp. 4709–4721.
[19] Y. Chen, W. Yu and T. Pock, *On learning optimized reaction diffusion processes for effective image restoration*, in IEEE CVPR, 2015, pp. 87–90.
[20] N. Dey, L. Blanc-Féraud, Z. K. C. Zimmer, J. C. Olivo-Marin and J. Zerubia, *A deconvolution method for confocal microscopy with total variation regularization*, in IEEE Intern. Symp. on Biomedical Imaging: Macro to Nano, 2004, pp. 1223–1226.
[21] E. Esser, X. Zhang and T. F. Chan, *A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science*, SIAM J. Imaging Sci., 3 (2010), pp. 1015–1046.
[22] T. Goldstein and S. Osher, *The split bregman method for l1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
[23] Y. Y. Haimes, L. S. Lasdon and D. A. Wismer, *On a bicriterion formulation of the problems of integrated system identification and system optimization*, IEEE Trans. Syst. Man Cyber., 1 (1971), pp. 296–297.
[24] M. Hintermüller, K. Ito and K. Kunisch, *The primal-dual active set strategy as a semismooth Newtons method*, SIAM J. Optim., 13 (2003), pp. 865–888.
[25] D. Krishnan, P. Lin and A. M. Yip, *A primal-dual active-set method for non-negativity constrained total variation deblurring problems*, IEEE Trans. Image Process., 16 (2007), pp. 2766–2777.
[26] D. Krishnan, Q. V. Pham and A. M. Yip, *A primal dual active set algorithm for bilaterally constrained total variation deblurring and piecewise constant Mumford-Shah segmentation problems*, Adv. Comput. Math., 31 (2009), pp. 237–266.
[27] E. C. Larson and D. M. Chandler, *Most apparent distortion: full-reference image quality assessment and the role of strategy*, J. Electron. Imaging, 19 (2010), 011006.
[28] Y. N. Law, H. K. Lee and A. M. Yip, *A multi-resolution stochastic level set method for Mumford-Shah image segmentation*, IEEE Trans. Image Process., 17 (2008), pp. 2289–2300.
[29] R. Liu, Z. Lin, W. Zhang and Z. Su, *Learning PDEs for Image Restoration via Optimal Control*, in ECCV, 2010, pp. 115–128.
[30] D. Martin, C. Fowlkes, D. Tal and J. Malik, *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics*, in IEEE ICCV, 2001.
[31] R. M. Mersereau and R. W. Schafer, *Comparative study of iterative deconvolution algorithms*, in IEEE ICASSP, 1978, pp. 192–195.
[32] J. M. Morel and S. Solimini, Variational Methods in Image Segmentation, Birkhauser Boston Inc., Cambridge, USA, 1995.
[33] N. Paragios, Y. Chen and O. D. Faugeras, Handbook of Mathematical Models in Computer Vision, Springer Science and Business Media, 2006.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press

[34] M. Persson, D. Bone and H. Elmqvist, *Total variation norm for three-dimension iterative reconstruction in limited view angle tomography*, Phys. Med. Biol., 46 (2001), pp. 853-866.
[35] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti et al., *Image database TID2013: peculiarities, results and perspectives*, Signal Process. Image Commun., 30 (2015), pp. 57–77.
[36] L. I. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[37] R. W. Schafer, R. M. Mersereau and M. A. Richards, *Constrained iterative restoration algorithms*, Proc. IEEE, 69 (1981), pp. 432–450.
[38] O. Scherzer, Handbook of Mathematical Methods in Imaging, Springer-Verlag, New York, USA, 2011.
[39] H. R. Sheikh, M. F. Sabir and A. C. Bovik, *A statistical evaluation of recent full reference image quality assessment algorithms*, IEEE Trans. Image Process., 15 (2006), pp. 3440–3451.
[40] C. Tai, X. Zhang and Z. Shen, *Wavelet frame based multiphase image segmentation*, SIAM J. Imaging Sci., 6 (2013), pp. 2521–2546.
[41] X. C. Tai, K. A. Lie, T. F. Chan and S. Osher, Image Processing Based on Partial Differential Equations, Springer-Verlag, New York, USA, 2007.
[42] X. C. Tai, K. Mørken, M. Lysaker and K. A. Lie, Scale Space and Variational Methods in Computer Vision, Springer Berlin Heidelberg, 2009.
[43] A. N. Tikhonov and V. Y. Arsenin, Solutions of Ill-Posed Problems, Winston and Sons, Washington, D. C., 1977.
[44] C. R. Vogel, Computational Methods for Inverse Problems, SIAM, Philadelphia, USA, 2002.
[45] J. Weickert, S. Ishikawa and A. Imiya, *Linear scale-space has first been proposed in Japan*, J. Math. Imaging Vis., 10 (1999), pp. 237–252.
[46] C. Wu and X. C. Tai, *Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models*, SIAM J. Imaging Sci., 3 (2010), pp. 300–339.
[47] Y. L. You and M. Kaveh, *Blind image restoration by anisotropic regularization*, IEEE Trans. Image Process., 8 (1999), pp. 396–407.

https://doi.org/10.4208/nmtma.2017.m1653 Published online by Cambridge University Press