Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2023-0091

Vol. **xx**, No. x, pp. 1-35
x 2024

Fusing Infrared and Visible Images via a First-Order Model

Wenli Yang[1], Zhongyi Huang[2,*] and Wei Zhu[3]

[1] *School of Mathematics, China University of Mining and Technology, Xuzhou 221116, China*
[2] *Department of Mathematics, Tsinghua University, Beijing 100084, China*
[3] *Department of Mathematics, University of Alabama, Tuscaloosa, AL 35487, USA*

Received 1 August 2023; Accepted (in revised version) 14 November 2023

Abstract. We propose a novel first-order non-convex model for the fusion of infrared and visible images. It maintains thermal radiation information by ensuring that the fused image has similar pixel intensities as the infrared image, and it preserves the appearance information, including the edges and texture of the source images, by enforcing similar gray gradients and pixel intensities as the visible image. Our model could effectively reduce the staircase effect and enhance the preservation of sharp edges. The maximum-minimum principle of the model with Neumann boundary condition is discussed and the existence of a minimizer of our model in $W^{1,2}(\Omega)$ is also proved. We employ the augmented Lagrangian method (ALM) to design a fast algorithm to minimize the proposed model and establish the convergence analysis of the proposed algorithm. Numerical experiments are conducted to showcase the distinctive features of the model and to provide a comparison with other image fusion techniques.

AMS subject classifications: 65M32, 94A08, 65K10
Key words: Image fusion, variational model, augmented Lagrangian methods.

1. Introduction

Image fusion refers to the process of acquiring the same scene from multiple source channels and integrating complementary multi-focus, multi-modal, multi-temporal, and/or multi-viewpoint images into a new image. This enhances its suitability for human or machine perception compared to the individual source images. Image fusion techniques can be classified into five categories: multi-view image fusion, multi-modal

*Corresponding author. *Email addresses:* yangwl19@cumt.edu.cn (W. Yang), zhongyih@mail.tsinghua.edu.cn (Z. Huang), wzhu7@ua.edu (W. Zhu)

image fusion, multi-temporal image fusion, multi-focus image fusion, and image fusion for image restoration. Infrared and visible image fusion, as a crucial and indispensable branch in the field of image fusion, falls under the category of multi-modal image fusion. It holds significant significance in night vision technology, security monitoring and image dehazing. For example, Zhu *et al.* [55] proposed a novel fast single image dehazing algorithm based on artificial multiexposure image fusion, which first combines the global and local details of the gamma-corrected images by a pixelwise weight computation, and then balances both image luminance and color saturation, finally can obtain high-visibility images by the effective and efficient mitigation of adverse haze effects. Image fusion can be also performed at pixel, feature, and symbol levels [25]. Infrared and visible image fusion is categorized under pixel-level image fusion. Prior to pixel-level fusion, it is essential to perform multi-sensor image registration. In this paper, we assume that all source images have been registered.

Over the past few decades, several techniques have been proposed for pixel-level fusion. These include the Laplacian pyramid (LP) [5,41,43], the discrete wavelet transform (DWT) [11], the dual-tree complex wavelet transform (DTCWT) [20, 21], the curvelet transform (CVT) [16, 33], the non-subsampled contourlet transform (NSCT) theory [13, 14], the multi-resolution singular value decomposition (MSVD) [32], guided filtering fusion (GFF) [24], autoencoder-based approaches [15], and other techniques [38]. Recently, deep learning-based fusion methods have also been developed, including the DenseFuse method [22], the RFN-Nest method [23], the SDNet method [49], the SeAFusion method [40], image fusion based on proportional maintenance of gradient and intensity (PMGI) [50], image fusion based on convolutional neural network (IFCNN) [17, 51], and fusion method based on generative adversarial networks (FusionGAN) [26].

In 2016, Ma *et al.* [25, 27] formulated the problem of fusing infrared and visible images by minimizing the following objective function:

$$F(s) = \frac{1}{2}\|s-u\|_2^2 + \lambda\|\nabla s - \nabla v\|_1,$$

where the first term constrains the fused image s to have similar pixel intensities with the infrared image u, the second term requires that the fused image s and the visible image v have similar gradients, and λ is a positive parameter controlling the trade-off between the two terms.

Notice that in the above functional, the second term uses the total variation, which could help s keep the edge locations as v. However, as discussed in [2,30,35], this total variation based regularizer could give rise to the staircase effect and the loss of image contrast. To remedy these unfavorable features, especially for the staircase effect, many higher-order variational models have been developed in the literature. These models employ different regularizers like total generalized variation, Euler's elastica, nonlinear fourth-order diffusive term, and second-order derivatives [4, 7, 39, 47]. Even though these higher-order models have proven effective for reducing the staircase effect, these higher-order models are intractable both analytically and numerically. To avoid the use

of higher-order derivatives, an interesting approach is to modify the potential function of the regularizer [42, 52]. This involves applying the L^p-norm of the image gradient with $p > 1$ for regions with relatively small image gradients, while imposing the original total variation on regions with relatively large image gradients. This modified regularizer successfully suppresses the staircase effect. This type of regularizer can be traced back to Huber [18] and is widely used in the context of robust statistics. The potential function used in the Huber model [18] is defined as follows:

$$\phi_a(x) = \begin{cases} \dfrac{1}{2a}x^2, & |x| \le a, \\ |x| - \dfrac{1}{2}a, & |x| > a, \end{cases}$$

where $a > 0$ is a parameter. It has also been extensively studied in image processing by Mila Nikolova and her collaborators [34]. Although the Huber model can effectively eliminate the staircase effect, it does not address the issue of blurred boundaries. In 2021, Zhu [53] proposed a novel first-order variational model for image restoration, defined by the following functional:

$$E(u) = \lambda \int_\Omega \phi_a(|\nabla u|) + \frac{\lambda}{2}\int_\Omega (f - K * u)^2,$$

where K represents a blurring operator, the new potential function ϕ_a reads

$$\phi_a(x) = \begin{cases} \dfrac{1}{2a}x^2, & |x| \le a, \\ a\ln|x| + \dfrac{a}{2} - a\ln a, & |x| > a, \end{cases} \tag{1.1}$$

and $a > 0$ is a parameter. The aim was to avoid the staircase effect while preserving image contrasts. Specifically, this regularizer has a lower growth rate than total variation (or even any power function $|x|^r$ with $r > 0$) for regions with relatively large image gradients to maintain image contrast, while imposing the L^2-norm of the gradient for regions with relatively small image gradients to reduce the staircase effect. The experiments in [53] demonstrated that this model efficiently removes noise while preserving image contrast and suppressing the staircase effect. Importantly, when applied to deblurring problems, this model generates deblurred images with cleaner boundaries than the well-known ROF model [36].

However, due to the non-convexity of the functional $E(u)$, the existence of its minimizer is still missing. In this paper, we propose modifying the potential function ϕ_a by adding a quadratic function with a small coefficient to overcome the above existence issue, while maintaining the specific features of the original regularizer. With this new regularizer, we develop a novel first-order variational model for the infrared and visible image fuse problem.

In the literature, numerous fast algorithms have been introduced for total variation-based models [29,36]. These include the primal-dual method [8], Chambolle's method [6], the split Bregman method [35], the augmented Lagrangian method (ALM) [9, 46,

54], alternating direction method of multipliers (ADMM) [44], and inertial proximal alternating direction method of multipliers (IPADMM) [10]. In our work, to minimize our proposed functional, we plan to develop an ALM that transforms it into the search for a saddle point of an augmented Lagrangian functional. This can be accomplished by iteratively and alternately minimizing several comparatively simpler functionals. Typically, these resulting functionals either have closed-form solutions or can be solved efficiently using fast solvers, such as fast Fourier Transforms (FFTs).

As a summary, the main contribution of this paper lies in the following three aspects:

- We propose a novel first-order non-convex model for the fusion of infrared and visible images, which could effectively eliminate noise, reduce the staircase effect and preserve sharp edges.
- The existence of a minimizer of our non-convex model in $W^{1,2}(\Omega)$ is proved.
- The maximum-minimum principle of our model with Neumann boundary condition is discussed.

The organization of this paper is as follows. In Section 2, we first present a novel first-order model for image fusion. We give the proof of the existence of the minimizer of our model in $W^{1,2}(\Omega)$. In Section 3, we provide detailed explanations of the development of the ALM algorithm. In Section 4, we present the results of numerical experiments conducted to validate the features of the proposed model, particularly the reduction of the staircase effect and the preservation of image contrast. Finally, the conclusion is given in Section 5.

2. A first-order image fusion model

Let $u, v,$ and s denote the infrared, visible and fused images respectively. Our objective is to combine the signal information from the infrared and visible images to create a new image. The fused image should encompass the detailed signals such as textures and edges from the visible image, as well as include the thermal radiation information that is absent in the visible image due to lighting conditions but captured by the infrared image.

Table 1: The meaning of the variables.

Variable	Meaning
u	the infrared image
v	the visible image
s	the fused image
q	$q = s - v$, the grayscale difference between the fused image and the visible image
$\boldsymbol{p}$	$\boldsymbol{p} = \nabla q$, the gradient of q

In general, since the thermal radiation is characterized by pixel intensities, it is expected that the pixel intensities of s and u are similar. To quantify the difference, we introduce the following empirical error:

$$E_1(s) = \frac{1}{2}\int_\Omega (s-u)^2 dx.$$

According to the human visual system, rich edge and texture feature information is observed when the spatial frequency is large. The spatial frequency is determined based on the distribution of gradients and pixel intensities in the image. Consequently, the essential features such as edges and textures in image v are represented by the gray gradients and pixel intensities. Therefore, in order to utilize the detailed appearance information in image v, it is necessary for the gray gradients and pixel intensities of images s and v to be close. In this work, we propose the following functional to quantify the difference between the two images s and v:

$$E_2(s) = \int_\Omega \left[\phi_a(|\nabla s - \nabla v|) + \eta|\nabla s - \nabla v|^2\right] dx + \frac{\lambda}{2}\int_\Omega (s-v)^2 dx,$$

where λ, η are positive parameters, and the potential function $\phi_a(x)$ is defined as (1.1). This choice of regularizer could suppress the staircase effect effectively and largely preserve image contrast. More importantly, as shown latter, the term $\eta|\nabla s - \nabla v|^2$ ensures the existence of solutions.

By combining the above two functionals $E_1(s)$ and $E_2(s)$, we propose the following first-order variational model for the image fusion problem:

$$s^* = \operatorname*{argmin}_s E(s) = \operatorname*{argmin}_s \big\{E_2(s) + \alpha E_1(s)\big\},$$

where

$$E(s) = \int_\Omega \left[\phi_a(|\nabla(s-v)|) + \eta|\nabla(s-v)|^2\right] dx + \frac{\lambda}{2}\int_\Omega (s-v)^2 dx + \frac{\alpha}{2}\int_\Omega (s-u)^2 dx. \quad (2.1)$$

Note that this model incorporates information from both the given infrared image u and the visible image v. $\lambda, \alpha > 0$ are tuning parameters that balance the regularization term and the fitting terms. The larger λ (or α) is, the more information in the visible image v (or the infrared image u) will be contained in the fused image s. η is a small positive parameter and it is used to ensure the existence of a minimizer of our model in $W^{1,2}(\Omega)$. η could be arbitrarily small, and the smaller it is, the better the image contrast will be preserved.

In our model, the parameter $a > 0$ is used to define regions with small or large magnitude of the gradient of the difference $s-v$. On one hand, in the region with $|\nabla(s-v)| \le a$, Tikhonov regularization is employed to impose a higher level of regularity on $s-v$ compared to total variation, thereby reducing the staircase effect. On the other hand, in the region where $|\nabla(s-v)| > a$, the regularizer has a lower growth rate

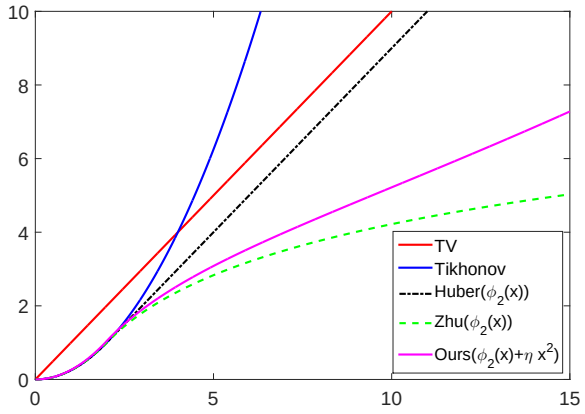

Figure 1: The graphs of different potential functions.

than total variation, which weakens the competition of the regularization term with the fitting term and thus helps to maintain image contrast.

In Fig. 1, we present the graphs of potential functions for the total variation regularization term, Tikhonov regularization term, Huber regularization term, the regularization term in [53] with $a = 2$, and our regularization term with $a = 2, \eta = 0.01$.

To illustrate the impact of the new regularizer on the image fusion results, we compare the performance of our model with models that use total variation regularization and Huber regularization in Section 4. For convenience, we will refer to these corresponding models as the TV model and the Huber model, respectively.

In fact, if we define $q = s - v$, the problem (2.1) can be reformulated as follows:

$$q^* = \underset{q}{\operatorname{argmin}}\, E(q), \tag{2.2}$$

where

$$E(q) = \int_\Omega \left[\phi_a(|\nabla q|) + \eta|\nabla q|^2\right] dx + \frac{\lambda}{2}\int_\Omega q^2 dx + \frac{\alpha}{2}\int_\Omega \big(q - (u - v)\big)^2 dx. \tag{2.3}$$

Once we find the optimal solution q^*, the fused image s^* can be determined using the relation $s^* = q^* + v$.

It is often seen that some hard constraints are imposed on the objective functional, such as the characteristic (indication) function $\mathcal{I}_{\mathbf{C}}(\cdot)$ [1], defined as $\mathcal{I}_{\mathbf{C}}(u) = 0$ if $u \in \mathbf{C}$ and $\mathcal{I}_{\mathbf{C}}(u) = \infty$ if $u \notin \mathbf{C}$, where $\mathbf{C}$ is a given set. If the solution satisfies the maximum-minimum principle, there is no need to add hard constraints to the objective functional. Next, we discuss the maximum-minimum principle to the problem (2.5) by utilizing Stampacchia's truncation method [19, 45].

The Gateaux derivative of E at q is

$$E'(q) = -\text{div}\left(\phi_a'(|\nabla q|)\frac{\nabla q}{|\nabla q|} + 2\eta\nabla q\right) + \lambda q + \alpha\big(q - (u - v)\big). \tag{2.4}$$

Let us consider the following nonlinear parabolic problem in 2D case:

$$\begin{cases} \dfrac{\partial q}{\partial t} = \operatorname{div}\left(\phi_a'(|\nabla q|)\dfrac{\nabla q}{|\nabla q|} + 2\eta\nabla q\right) - \lambda q - \alpha\big(q-(u-v)\big) & \text{in } \Omega\times(0,\infty), \\ \dfrac{\partial q}{\partial \vec{n}} = 0 & \text{on } \partial\Omega\times(0,\infty), \\ q = \dfrac{\alpha}{\lambda+\alpha}(u-v) & \text{on } \Omega\times\{t=0\}. \end{cases} \tag{2.5}$$

Theorem 2.1. *Consider the problem* (2.5) *with any bounded measurable function* u *and* v*, then the solution* q *of* (2.5) *verifies*

$$\frac{\alpha}{\lambda+\alpha}\inf_{x\in\Omega}\{u(x)-v(x)\} \le q(x,t) \le \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\} \quad \textit{on } \Omega\times(0,\infty).$$

Proof. For the problem (2.5), we only prove the maximum principle. The minimum principle follows from the reverse contrast invariance and the maximum principle when applied to the initial data $-u$ and $-v$.

Let $G\in C^1(R)$ be a function with $G(l)=0$ on $(-\infty,0]$ and $0<G'(l)\le C$ on $(0,\infty)$ for some constant C. Now we define

$$\begin{aligned} H(l) &:= \int_0^l G(s)ds, \quad l\in R, \\ J(t) &:= \int_\Omega H\left(q(x,t) - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)dx, \quad t\in[0,\infty). \end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} &\int_\Omega \left|G\left(q(x,t) - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)\frac{\partial q}{\partial t}(x,t)\right|dx \\ &\le C\left\|q(t) - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right\|_{L^2(\Omega)}\left\|\frac{\partial q}{\partial t}(t)\right\|_{L^2(\Omega)}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial J}{\partial t} &= \int_\Omega G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)\frac{\partial q}{\partial t}dx \\ &= \int_\Omega G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)\operatorname{div}\left(\phi_a'(|\nabla q|)\frac{\nabla q}{|\nabla q|} + 2\eta\nabla q\right)dx \\ &\quad + \int_\Omega G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)\big(\alpha(u-v) - (\lambda+\alpha)q\big)dx \\ &= \int_{\partial\Omega} G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)\underbrace{\left\langle \phi_a'(|\nabla q|)\frac{\nabla q}{|\nabla q|} + 2\eta\nabla q, n\right\rangle}_{=0}dS \end{aligned}$$

$$
\begin{aligned}
&- \int_\Omega \underbrace{G'\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)}_{\geq 0} \underbrace{\left\langle \nabla q, \phi_a'(|\nabla q|)\frac{\nabla q}{|\nabla q|} + 2\eta\nabla q\right\rangle}_{\geq 0} dx \\
&+ \int_{\Omega\bigcap\{q-\frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\leq 0\}} \big(\alpha(u-v)-(\lambda+\alpha)q\big) \\
&\qquad\qquad \times \underbrace{G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)}_{=0} dx \\
&+ \int_{\Omega\bigcap\{q-\frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}>0\}} \underbrace{\big(\alpha(u-v)-(\lambda+\alpha)q\big)}_{\leq 0} \\
&\qquad\qquad \times \underbrace{G\left(q - \frac{\alpha}{\lambda+\alpha}\sup_{x\in\Omega}\{u(x)-v(x)\}\right)}_{\geq 0} dx \leq 0.
\end{aligned}
$$

We also have $G(l) \leq Cl$, then $H(l) \leq Cl^2/2$. Therefore,

$$
\begin{aligned}
0 \leq J(t) &\leq \int_\Omega H\left(q(x,t) - \frac{\alpha}{\lambda+\alpha}\big(u(x)-v(x)\big)\right) dx \\
&\leq \frac{C}{2}\left\| q(t) - \frac{\alpha}{\lambda+\alpha}\big(u(x)-v(x)\big)\right\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Since $q \in C([0,T];L^2(\Omega))$, we have $J(0)=0$, which proves the continuity of $J(t)$ in 0. According to $J \in C([0,\infty))$, $J(0)=0$, $J(t)\geq 0$ on $[0,\infty)$ and $\partial J/\partial t \leq 0$, we get $J \equiv 0$ on $[0,\infty)$. Hence, $q(x,t) \leq (\alpha/(\lambda+\alpha))\sup_{x\in\Omega}\{u(x)-v(x)\}$ on $\Omega\times(0,\infty)$. □

Corollary 2.1. *Consider the problem* (2.5) *with any bounded measurable function* u *and* v*, and*

$$
\operatorname*{ess\,inf}_{x\in\Omega} u(x) = 0, \quad \operatorname*{ess\,sup}_{x\in\Omega} u(x) = 1, \quad \operatorname*{ess\,inf}_{x\in\Omega} v(x) = 0, \quad \operatorname*{ess\,sup}_{x\in\Omega} v(x) = 1.
$$

Then the solution q *of* (2.5) *verifies*

$$
-\frac{\alpha}{\lambda+\alpha} \leq q(x,t) \leq \frac{\alpha}{\lambda+\alpha} \quad \text{on } \Omega\times(0,\infty).
$$

We first study the existence of the minimizer of the functional $E(q)$ in one dimensional with $\Omega=[l_1,l_2]$. At this time, we can deduce that when $\eta>0$ is sufficient small, the minimizer q of $E(q)$ is allowed to have large jumps, thereby maintaining image contrast.

Lemma 2.1. *Consider the function* $\psi(x)=\phi_a(x)+\eta x^2$ *over* $\mathbb{R}$*, where* $\phi_a(x)$ *is defined as in* (1.1)*, and* $\eta>0$*. If* $\eta \geq 1/(2a)$*, then* $\psi(x)$ *is convex on* $\mathbb{R}$*.*

Proof. Substituting the expression of $\phi_a(x)$ into the function $\psi(x)$, we get

$$\psi(x) = \begin{cases} \dfrac{1}{2a}x^2 + \eta x^2, & |x| \le a, \\ a\ln|x| + \dfrac{a}{2} - a\ln a + \eta x^2, & |x| > a. \end{cases}$$

Its second derivative is

$$\psi''(x) = \begin{cases} \dfrac{1}{a} + 2\eta, & |x| \le a, \\ -\dfrac{a}{x^2} + 2\eta, & |x| > a. \end{cases}$$

So, if $\eta \ge 1/(2a)$, then $\psi''(x) \ge 0$. Since $\psi(x)$ is $\mathbf{C}^1(\mathbb{R})$ and $\psi'(x)$ is non-decreasing on $\mathbb{R}$, $\psi(x)$ is convex on $\mathbb{R}$. □

As mentioned before, the parameter a is a cut-off point that divides the image into two regions, over which different regularization terms are imposed. If the gray value of the visible image and infrared image falls within the interval $[0, 1]$ and the mesh size is chosen as $h = 0.01$, then the range of the grayscale gradient will be $[0, 100]$, and the fusion effect is best when a is taken as 2. The more interesting and important case is when the parameters η and a satisfy the inequality $0 < \eta < 1/(2a)$, since it allows sharp transitions of $q(x)$, where $q(x)$ is a minimizer of $E(q)$.

Theorem 2.2. *If $u(x)$ and $v(x)$ are the simple bounded measurable functions, and $\eta > 0$, then $E(q)$ has a minimizer in $H^1(\Omega)$.*

Proof. Let $\psi_\eta(|q'|) = \phi_a(|q'|) + \eta(q')^2$, and $\psi_\eta^{**}(|q'|)$ be the lower convex envelope of $\psi_\eta(|q'|)$, just like Fig. 2. If $\eta \ge 1/(2a)$, then $\psi_\eta^{**}(|q'|) = \psi_\eta(|q'|)$. At this time, $E(q)$ has a minimizer in $H^1(\Omega)$. If $\eta < 1/(2a)$, then

$$\psi_\eta^{**}(|q'|) = \begin{cases} \psi_\eta(|q'|), & |q'| \in [0, \theta_1] \cup [\theta_2, +\infty), \\ \left(\dfrac{1}{a} + 2\eta\right)\theta_1|q'| - \left(\dfrac{1}{2a} + \eta\right)\theta_1^2, & |q'| \in (\theta_1, \theta_2), \end{cases}$$

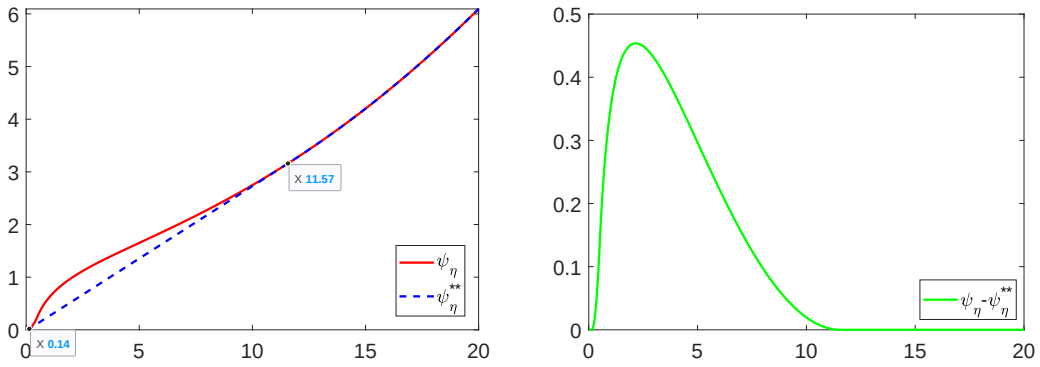

Figure 2: Left: the graph of the ψ_η and ψ_η^{**} when $a = 0.5$ and $\eta = 0.01$, in which case $\theta_1 = 0.14$ and $\theta_2 = 11.57$, right: the error $\psi_\eta - \psi_\eta^{**}$.

where $\theta_1 > 0$ and $\theta_2 > 0$ satisfy

$$\begin{aligned} &\frac{\theta_1}{a} + 2\eta\theta_1 = \frac{a}{\theta_2} + 2\eta\theta_2, \\ &a\ln|\theta_2| + \frac{a}{2} - a\ln a + \eta\theta_2^2 = -\left(\frac{1}{2a} + \eta\right)\theta_1^2 + \left(\frac{1}{a} + 2\eta\right)\theta_1\theta_2, \end{aligned} \tag{2.6}$$

and

$$\frac{d\psi_\eta^{**}(|q'|)}{dq'} = \begin{cases} \dfrac{q'}{a} + 2\eta q', & |q'| \le \theta_1, \\ \dfrac{a}{q'} + 2\eta q', & |q'| \ge \theta_2, \\ \left(\dfrac{1}{a} + 2\eta\right)\theta_1, & \theta_1 < q' < \theta_2, \\ -\left(\dfrac{1}{a} + 2\eta\right)\theta_1, & -\theta_2 < q' < -\theta_1. \end{cases}$$

Then using [12, Theorem 9.1], we have

$$\min_{q\in H^1(\Omega)} E^{**}(q) = \inf_{q\in H^1(\Omega)} E(q),$$

where

$$E^{**}(q) = \int_\Omega \psi_\eta^{**}(|q'|)dx + \frac{\lambda}{2}\int_\Omega q^2 dx + \frac{\alpha}{2}\int_\Omega \big(q - (u-v)\big)^2 dx.$$

Since u and v are the simple bounded measurable functions, then we have the following partition of Ω:

$$\Omega = (\cup_j S_j) \cup \Sigma \quad \text{with} \quad |\Sigma| = 0,$$

and

$$u(x) - v(x) = c_j, \quad \forall x \in S_j,$$

where c_j are constants. Suppose $q_0(x) \in H^1(\Omega)$ such that

$$E^{**}(q_0) = \inf_{q\in H^1(\Omega)} E(q),$$

then $q_0(x)$ satisfies the following problem:

$$\begin{aligned} &\frac{d}{dx}\left[\frac{d}{dq_0'}\psi_\eta^{**}(|q_0'|)\right] = \lambda q_0(x) + \alpha\big(q_0(x) - (u(x) - v(x))\big), \quad \text{a.e. in } \Omega, \\ &\frac{d}{dq_0'}\psi_\eta^{**}(|q_0'|)|_{x=l_1} = \frac{d}{dq_0'}\psi_\eta^{**}(|q_0'|)|_{x=l_2} = 0. \end{aligned}$$

And therefore we have $\forall x \in \Omega$,

$$\frac{d}{dq_0'}\psi_\eta^{**}(|q_0'(x)|) = \int_{l_1}^x \big[\lambda q_0(t) + \alpha(q_0(t) - (u(t) - v(t)))\big]dt. \tag{2.7}$$

Now we introduce the following notation:

$$\Omega_1 = \left\{ x \in \Omega : \frac{d\psi_\eta^{**}(|q_0'(x)|)}{dq_0'} = \left(\frac{1}{a} + 2\eta\right)\theta_1 \right\},$$
$$\Omega_2 = \left\{ x \in \Omega : \frac{d\psi_\eta^{**}(|q_0'(x)|)}{dq_0'} = -\left(\frac{1}{a} + 2\eta\right)\theta_1 \right\}.$$

To show that $q_0(x)$ is also a minimizer of $E(q)$, it is enough to show that $|\Omega_1| = 0$ and $|\Omega_2| = 0$. By contradiction, if $|\Omega_1| > 0$, then there is at least one measurable set S_j such that $|\Omega_1 \cap S_j| > 0$. Let $S_j^* = \Omega_1 \cap S_j$, we have

$$\left(\frac{1}{a} + 2\eta\right)\theta_1 = \int_{l_1}^{x} \big[\lambda q_0 + \alpha(q_0 - (u - v))\big]dt, \quad \forall x \in S_j^*.$$

When derivating with respect to x on both sides of the above equality, we find

$$q_0(x) = \frac{\alpha}{\lambda + \alpha}\big(u(x) - v(x)\big) = \frac{\alpha}{\lambda + \alpha}c_j, \quad \text{a.e. in } S_j^*.$$

Applying the Stampacchia's result [19], one obtains

$$q_0'(x) = 0, \quad \text{a.e. in } S_j^*.$$

But this is a contradiction, according to $\theta_1 < q_0'(x) < \theta_2$, $\forall x \in S_j^* \subseteq \Omega_1$. In the same way one can prove $|\Omega_2| = 0$. □

From Theorem 2.1 and Eq. (2.7), we can obtain the inequality

$$\begin{aligned}
2\eta|q_0'(x)| &\le \left|\frac{d}{dq_0'}\psi_\eta^{**}(|q_0'(x)|)\right| \\
&= \left|\int_{l_1}^{x} \big[\lambda q_0(t) + \alpha(q_0(t) - (u(t) - v(t)))\big]dt\right| \\
&= \left|\int_{l_1}^{x} \big[(\lambda + \alpha)q_0(t) - \alpha(u(t) - v(t))\big]dt\right| \\
&\le \left|\int_{l_1}^{x} \left[(\lambda + \alpha)\frac{\alpha}{\lambda + \alpha}\sup_{x\in\Omega}\big\{u(t) - v(t)\big\} - \alpha\inf_{x\in\Omega}\big\{u(t) - v(t)\big\}\right]dt\right| \\
&\le \alpha|\Omega|\Big(\sup_{x\in\Omega}\big\{u(t) - v(t)\big\} - \inf_{x\in\Omega}\big\{u(x) - v(x)\big\}\Big).
\end{aligned}$$

So, $q_0(x)$ has the properties

$$|q_0'(x)| \le \frac{\alpha}{2\eta}|\Omega|\Big(\sup_{x\in\Omega}\big\{u(t) - v(t)\big\} - \inf_{x\in\Omega}\big\{u(x) - v(x)\big\}\Big), \quad \text{a.e. in } \Omega. \tag{2.8}$$

When η is sufficient small, $|q_0'(x)|$ is allowed to be large, so $q_0(x)$ can have large jumps to maintain image contrast, where $q_0(x)$ is a minimizer of $E(q)$.

Next, we consider the existence of the solution to the problem

$$\inf_{q\in W^{1,2}(\Omega)} \int_\Omega \left[\phi_a(|Dq|) + \eta|Dq|^2\right] dx + \frac{\lambda}{2}\int_\Omega q^2 dx + \frac{\alpha}{2}\int_\Omega \big(q-(u-v)\big)^2 dx, \tag{2.9}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded, open, and connected domain with Lipschitz boundary, q is a scalar function that belongs to $W^{1,2}(\Omega)$, $W^{1,2}(\Omega) = H^1(\Omega)$ consists of all locally summable functions $q : \Omega \to \mathbb{R}$ such that for each multiindex $|\alpha| \le 1$, $D^\alpha q$ exists in the weak sense and belongs to $L^2(\Omega)$. We denote

$$\mathcal{F}(q) := \int_\Omega g(x, q, \nabla q)dx, \tag{2.10}$$

where

$$g(x, q, \nabla q) = \phi_a(|\nabla q|) + \eta|\nabla q|^2 + \frac{\lambda}{2}q^2 + \frac{\alpha}{2}\big(q-(u-v)\big)^2. \tag{2.11}$$

$g : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ is a Carathéodory function as in (2.11), that is, $g(\cdot, s, \xi) : \Omega \to \mathbb{R}$ is measurable for all $(s, \xi) \in \mathbb{R} \times \mathbb{R}^2$, $g(x, \cdot, \xi) : \mathbb{R} \to \mathbb{R}$ is measurable for all $(x, \xi) \in \Omega \times \mathbb{R}^2$, and $g(x, s, \cdot) : \mathbb{R}^2 \to \mathbb{R}$ is continuous for almost every $(x, s) \in \Omega \times \mathbb{R}$. In fact it has been proved that the convexity or the quasi-convexity of the integrand g with respect to the last component ξ is necessary to the weak lower semi-continuity of the given integral [28,31]. g is quasi-convex in the sense of Morrey [31], that is, for almost every $x_0 \in \Omega, q_0 \in \mathbb{R}, p_0 \in \mathbb{R}^2$, and any $\varphi \in W_0^{1,\kappa}(\Omega)$ there holds

$$\frac{1}{\mathcal{L}^n(\Omega)}\int_\Omega g\big(x_0, q_0, p_0 + D\varphi(x)\big)dx \ge g(x_0, q_0, p_0),$$

where $\kappa \ge 1$, $\mathcal{L}^n$ denotes the n-dimensional Lebesgue measure. We also recall that in the scalar case, quasi-convexity and ordinary convexity are equivalent. So $\mathcal{F}$ is not weakly sequentially lower semi-continuous. The direct method of the calculus of variations, which is used to establish the existence of minima for the given integral, relies on the lower semi-continuity of the integral. However, it is important to note that convexity is not a necessary condition for the existence of minima [28]. Borrowing the proof idea from Theorem 2.2, the first step in dealing with such problems is to apply the relaxation theorem. Instead of minimizing $\mathcal{F}$ directly, one computes the minimizers of a relaxed functional $\mathcal{RF}$, which is defined as the largest weakly lower semi-continuous functional lower than $\mathcal{F}$. Let $X \subset W^{1,2}(\Omega)$ be a closed and convex subset. The relaxation $\mathcal{RF}$ of $\mathcal{F} : X \to \mathbb{R} \cup \{\infty\}$ is defined as follows:

$$\mathcal{RF}(q) := \inf\Big\{ \liminf_k \mathcal{F}(q_k) : \{q_k\} \subset X \ \text{ and } \ q_k \rightharpoonup q \ \text{ in } \ W^{1,2}(\Omega)\Big\}.$$

Lemma 2.2. *Since $g : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}\bigcup\{\infty\}$ is normal and satisfies*

$$g(x, q, t) \ge \eta|t|^2,$$

then $\mathcal{RF}$ is weakly sequentially lower semi-continuous in $W^{1,2}(\Omega)$.

Proof. The proof is similar to the proof of [37, Lemma 5.4]. □

The convexification g^c with respect to the last variable is defined as the largest convex integrand below g. Using Carathéodory's theorem, it follows that

$$g^c(x,q,t) = \inf\left\{\sum_{k=1}^{n+1}\lambda_k g(x,q,t_k) : 0\le\lambda_k\le 1, \sum_{k=1}^{n+1}\lambda_k t_k = t\right\}.$$

Lemma 2.3. *Let $g:\Omega\times\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}\bigcup\{\infty\}$ be as in* (2.11)*. Then the convex hull g^c of g with respect to the last variable is*

$$g^c(x,q,t) = \begin{cases} \dfrac{1}{2a}t^2+\eta t^2+\dfrac{\lambda}{2}q^2+\dfrac{\alpha}{2}\big(q-(u-v)\big)^2, & |t|\le\theta_1,\\ \left(\dfrac{1}{a}+2\eta\right)\theta_1|t|-\left(\dfrac{1}{2a}+\eta\right)\theta_1^2+\dfrac{\lambda}{2}q^2+\dfrac{\alpha}{2}\big(q-(u-v)\big)^2, & \theta_1<|t|<\theta_2,\\ a\ln|t|+\dfrac{a}{2}-a\ln a+\eta t^2+\dfrac{\lambda}{2}q^2+\dfrac{\alpha}{2}\big(q-(u-v)\big)^2, & |t|\ge\theta_2, \end{cases}$$

where θ_1 and θ_2 satisfy (2.6).

An example of the graphs of g and its convex hull g^c is shown in Fig. 3.

Lemma 2.4. *Let $g:\Omega\times\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}\bigcup\{\infty\}$ be as in* (2.11)*. Then*

$$\mathcal{RF}(q) = \mathcal{F}^c(q) := \int_\Omega g^c(x,q,Dq)dx, \quad q\in W^{1,2}(\Omega).$$

Proof. Define

$$g_k(x,q,t) = \phi_a(|t|)+\eta|t|^2+\frac{\lambda}{2}q^2+\frac{\alpha}{2}\big(q-(u-v)\big)^2+\frac{|t|^3}{k},$$

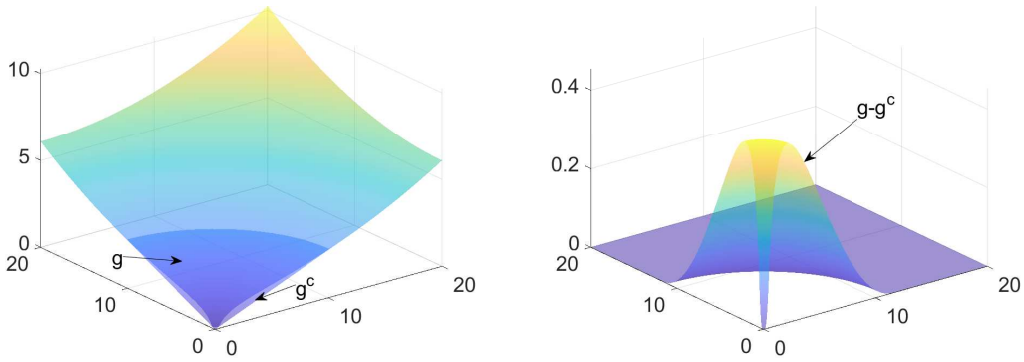

Figure 3: Left: the graph of the g and g^c when $a=0.5$, $\eta=0.01$ and all other terms independent of t are assumed to be 0, in which case $\theta_1=0.14$ and $\theta_2=11.57$, right: the error $g-g^c$.

then g_k pointwise converges to the function g satisfying

$$g_k(x,q,t) \geq \frac{|t|^3}{k}.$$

Moreover,

$$0 \leq g^c(x,q,t) \leq c\left(|q|^2+|t|^2\right), \quad (x,q,t) \in \Omega \times \mathbb{R} \times \mathbb{R}^2,$$

where

$$c = \max\left\{\frac{1}{2a}+\eta, \frac{\lambda+\alpha}{2}\right\} > 0.$$

Applying [37, Theorem 5.6], we have

$$\mathcal{RF}(q) = \mathcal{F}^c(q) := \int_\Omega g^c(x,q,Dq)dx, \quad q \in W^{1,2}(\Omega).$$

The proof is complete. □

Lemma 2.5. *Let $q \in W^{1,2}(\Omega)$, $u \in L^\infty(\Omega)$, $v \in L^\infty(\Omega)$, and assume that*

$$r \geq \operatorname*{ess\,sup}_{x\in\Omega}\{u(x)-v(x)\}, \quad s \leq \operatorname*{ess\,inf}_{x\in\Omega}\{u(x)-v(x)\}.$$

Then

$$\mathcal{F}^c(\min\{q,r\}) \leq \mathcal{F}^c(q), \quad \mathcal{F}^c(\max\{q,s\}) \leq \mathcal{F}^c(q).$$

Proof. We only show the first assertion, the second then follows from the first by considering $-q$ and $-(u-v)$. Denote $\tilde{q} = \min\{q,r\}$. Then

$$\begin{aligned}
&\int_\Omega g^c(x,\tilde{q},D\tilde{q})dx \\
&= \int_{\{q-r\leq 0\}} g^c(x,q,Dq)dx + \int_{\{q-r>0\}} \left(\frac{\lambda}{2}r^2 + \frac{\alpha}{2}\big(r-(u-v)\big)^2\right)dx \\
&\leq \int_{\{q-r\leq 0\}} g^c(x,q,Dq)dx + \int_{\{q-r>0\}} \left(\frac{1}{2a}t^2 + \eta t^2 + \frac{\lambda}{2}q^2 + \frac{\alpha}{2}\big(q-(u-v)\big)^2\right)dx \\
&\leq \int_\Omega g^c(x,q,Dq)dx.
\end{aligned}$$

The proof is complete. □

Denote

$$M := \left\{q \in W^{1,2}(\Omega) : \|q\|_\infty \leq \frac{\alpha}{\lambda+\alpha}\|u-v\|_\infty\right\}.$$

It follows from Lemma 2.5 that every minimizer of $\mathcal{F}^c$ over $W^{1,2}(\Omega)$ already lies in M. Thus minimizing $\mathcal{F}^c$ over $W^{1,2}(\Omega)$ is equivalent to minimizing $\mathcal{F}^c$ over M.

Theorem 2.3. *Suppose $\Omega \subset \mathbb{R}^2$ is bounded, open, connected, and Lipschitz, $g : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R} \bigcup \{\infty\}$ is defined as in* (2.11)*, $\mathcal{F} : W^{1,2}(\Omega) \to \mathbb{R} \bigcup \{\infty\}$ is given as in* (2.10)*, $u \in L^\infty(\Omega)$, and $v \in L^\infty(\Omega)$, then the functional $\mathcal{F}$ attains a minimizer in $W^{1,2}(\Omega)$.*

Proof. Let $\mathcal{F}^c : W^{1,2}(\Omega) \to \mathbb{R} \bigcup \{\infty\}$ be defined as in Lemma 2.4. From Lemmas 2.2 and 2.4, it follows that $\mathcal{F}^c$ is weakly sequentially lower semi-continuous on $W^{1,2}(\Omega)$. Denote

$$M := \left\{ q \in W^{1,2}(\Omega) : \|q\|_\infty \leq \frac{\alpha}{\lambda + \alpha} \|u - v\|_\infty \right\},$$

then it is sufficient to show that $\mathcal{F}^c|_M$ attains a minimizer. Now note that $\mathcal{F}^c(q) \geq \eta \|Dq\|_2^2$ for every $q \in W^{1,2}(\Omega)$. Recall that the t-level sets of a function $\mathcal{F} : X \to Y$ are defined as $\text{level}_t(\mathcal{F}) := \{x \in X : \mathcal{F}(x) \leq t\}, t \in Y$. Consequently, for every $t \in \mathbb{R}$, we have

$$\text{level}_t(\mathcal{F}^c|_M) \subset M \cap \left\{ q \in W^{1,2}(\Omega) : \|Dq\|_2^2 \leq \frac{t}{\eta} \right\}.$$

This shows that $\text{level}_t(\mathcal{F}^c|_M)$ is sequentially pre-compact, and thus $\mathcal{F}^c|_M$ is weakly sequentially coercive. Using [37, Theorem 5.1], the functional $\mathcal{F}^c$ attains a minimizer in $W^{1,2}(\Omega)$. According to [12, Theorem 9.1], we have

$$\inf_{q \in W^{1,2}(\Omega)} \mathcal{F}(q) = \inf_{q \in W^{1,2}(\Omega)} \mathcal{F}^c(q),$$

then the functional $\mathcal{F}$ attains a minimizer in $W^{1,2}(\Omega)$. □

3. Augmented Lagrangian method for the proposed model

To minimize our model (2.2), one may employ the associated gradient flow directly to find the solution. But it is often expensive, which is due to the constraint on stability conditions on the time step size.

In this work, we develop an augmented Lagrangian method for the minimization of our model. In fact, the augmented Lagrangian method has been widely used for minimizing non-differentiable or/and higher-order models [29, 39, 46, 54]. Precisely, for the minimization of functional Eq. (2.2), we propose an equivalent constrained optimization problem as follows:

$$\begin{aligned} &\min_{q, \boldsymbol{p}} \int_\Omega \left[\phi_a(|\boldsymbol{p}|) + \eta |\boldsymbol{p}|^2 \right] dx + \frac{\lambda}{2} \int_\Omega q^2 dx + \frac{\alpha}{2} \int_\Omega \big(q - (u - v)\big)^2 dx, \\ &\text{s.t.} \quad \boldsymbol{p} = \nabla q, \end{aligned}$$

and then consider the following augmented Lagrangian functional:

$$\begin{aligned} \mathcal{L}(q, \boldsymbol{p}; \boldsymbol{\lambda_1}) = &\int_\Omega \left[\phi_a(|\boldsymbol{p}|) + \eta |\boldsymbol{p}|^2 \right] dx + \frac{\lambda}{2} \int_\Omega q^2 dx + \frac{\alpha}{2} \int_\Omega \big(q - (u - v)\big)^2 dx \\ &+ \frac{\gamma_1}{2} \int_\Omega |\boldsymbol{p} - \nabla q|^2 dx + \int_\Omega \boldsymbol{\lambda_1} \cdot (\boldsymbol{p} - \nabla q) dx, \end{aligned} \tag{3.1}$$

where $\gamma_1 > 0$ is a penalization parameter to be chosen in numerical implementation, and $\boldsymbol{\lambda_1} \in \mathbb{R}^2$ is a Lagrange multiplier. Based on the theory of optimization, we need to find saddle points of $\mathcal{L}$ in order to find minimizers of the original functional $E(q)$. To find a saddle point of $\mathcal{L}$, we may apply an iterative algorithm: for each of q and $\boldsymbol{p}$, we fix the other one and seek a minimizer of the associated sub-problem, and then update the Lagrange multiplier once all q and $\boldsymbol{p}$ are advanced. The process will be repeated until the variables converge, which indicates the saddle point will be approximated. Therefore, we consider the minimization of the following two sub-problems:

$$\begin{aligned}\varepsilon_1(q;\boldsymbol{\lambda_1}) &= \frac{\lambda}{2}\int_\Omega q^2 dx + \frac{\alpha}{2}\int_\Omega \big(q-(u-v)\big)^2 dx + \frac{\gamma_1}{2}\int_\Omega (\boldsymbol{p}-\nabla q)^2 dx \\ &\quad + \int_\Omega \boldsymbol{\lambda_1}\cdot(\boldsymbol{p}-\nabla q)dx, \\ \varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1}) &= \int_\Omega \left[\phi_a(|\boldsymbol{p}|)+\eta|\boldsymbol{p}|^2\right]dx + \frac{\gamma_1}{2}\int_\Omega |\boldsymbol{p}-\nabla q|^2 dx + \int_\Omega \boldsymbol{\lambda_1}\cdot(\boldsymbol{p}-\nabla q)dx.\end{aligned}$$

For the subproblem of $\boldsymbol{p}$, the minimizer of $\varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1})$ has a closed-form solution. Note that it could be written as follows [52, 53]:

$$\varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1}) = \int_\Omega \phi_a(|\boldsymbol{p}|)dx + \Big(\frac{\gamma_1}{2}+\eta\Big)\int_\Omega |\boldsymbol{p}-\boldsymbol{p^*}|^2 dx + \tilde{C},$$

with $\boldsymbol{p^*} = (\gamma_1\nabla q - \boldsymbol{\lambda_1})/(\gamma_1+2\eta)$, and

$$\tilde{C} = \frac{\gamma_1}{2}|\nabla q|^2 - \boldsymbol{\lambda_1}\cdot\nabla q - \Big(\frac{\gamma_1}{2}+\eta\Big)|\boldsymbol{p^*}|^2$$

is independent of $\boldsymbol{p}$. We present the minimizer of $\varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1})$ by finding that of the integrand pointwisely.

Proposition 3.1. *The minimizer of $\varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1})$ can be determined as follows:*

$$\operatorname*{argmin}_{\boldsymbol{p}} \varepsilon_2(\boldsymbol{p};\boldsymbol{\lambda_1}) = \begin{cases} \dfrac{\gamma_1+2\eta}{1/a+\gamma_1+2\eta}\boldsymbol{p^*}, & |\boldsymbol{p^*}| \le a + \dfrac{1}{\gamma_1+2\eta}, \\ \dfrac{1+\sqrt{1-4a/((\gamma_1+2\eta)|\boldsymbol{p^*}|^2)}}{2}\boldsymbol{p^*}, & |\boldsymbol{p^*}| > a + \dfrac{1}{\gamma_1+2\eta}, \end{cases} \tag{3.2}$$

where $\boldsymbol{p^*} = (\gamma_1\nabla q - \boldsymbol{\lambda_1})/(\gamma_1+2\eta)$.

Note $\boldsymbol{p^*} = (p_1^*, p_2^*)$. By Proposition 3.1, one gets new updates of p_1 and p_2 as follows:

$$p_1^k(i,j) = \begin{cases} \dfrac{\gamma_1+2\eta}{1/a+\gamma_1+2\eta}p_1^*(i,j), & |\boldsymbol{p^*}(i,j)| \le a + \dfrac{1}{\gamma_1+2\eta}, \\ \dfrac{1+\sqrt{1-4a/((\gamma_1+2\eta)|\boldsymbol{p^*}(i,j)|^2)}}{2}p_1^*(i,j), & |\boldsymbol{p^*}(i,j)| > a + \dfrac{1}{\gamma_1+2\eta}, \end{cases}$$

$$p_2^k(i,j) = \begin{cases} \dfrac{\gamma_1 + 2\eta}{1/a + \gamma_1 + 2\eta} p_2^*(i,j), & |\boldsymbol{p}^*(i,j)| \le a + \dfrac{1}{\gamma_1 + 2\eta}, \\ \dfrac{1 + \sqrt{1 - 4a/((\gamma_1 + 2\eta)|\boldsymbol{p}^*(i,j)|^2)}}{2} p_2^*(i,j), & |\boldsymbol{p}^*(i,j)| > a + \dfrac{1}{\gamma_1 + 2\eta}, \end{cases}$$

where

$$p_1^*(i,j) = \frac{\gamma_1 \partial_1^+ q^{k-1}(i,j) - \lambda_{11}^{k-1}(i,j)}{\gamma_1 + 2\eta},$$
$$p_2^*(i,j) = \frac{\gamma_1 \partial_2^+ q^{k-1}(i,j) - \lambda_{12}^{k-1}(i,j)}{\gamma_1 + 2\eta},$$

and

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

The minimizer of $\varepsilon_1(q; \boldsymbol{\lambda_1})$ has no closed form, and it could be determined by the corresponding Euler-Lagrange equation as follows:

$$(\lambda + \alpha)q - \gamma_1 \triangle q = B,$$

where

$$B = \alpha(u - v) - \nabla \cdot (\gamma_1 \boldsymbol{p} + \boldsymbol{\lambda_1}).$$

Let h be the mesh size, $\Omega = \{(i,j) | 1 \le i \le M, 1 \le j \le N\}$ be the discretized image domain and each point (i,j) is a grid point. We define the discrete backward and forward differential operators with Neumann boundary condition as follows:

$$\partial_1^- q(i,j) = \begin{cases} \dfrac{q(i,j) - q(i-1,j)}{h}, & 1 < i \le M, \\ 0, & i = 1, \end{cases}$$
$$\partial_1^+ q(i,j) = \begin{cases} \dfrac{q(i+1,j) - q(i,j)}{h}, & 1 \le i < M, \\ 0, & i = M, \end{cases}$$
$$\partial_2^- q(i,j) = \begin{cases} \dfrac{q(i,j) - q(i,j-1)}{h}, & 1 < j \le N, \\ 0, & j = 1, \end{cases}$$
$$\partial_2^+ q(i,j) = \begin{cases} \dfrac{q(i,j+1) - q(i,j)}{h}, & 1 \le j < N, \\ 0, & j = N. \end{cases}$$

The gradient operators and the Laplace operator are defined accordingly

$$\nabla^{\pm} q(i,j) = \big(\partial_1^{\pm} q(i,j), \partial_2^{\pm} q(i,j)\big),$$
$$\triangle q(i,j) = \partial_1^+ \partial_1^- q(i,j) + \partial_2^+ \partial_2^- q(i,j).$$

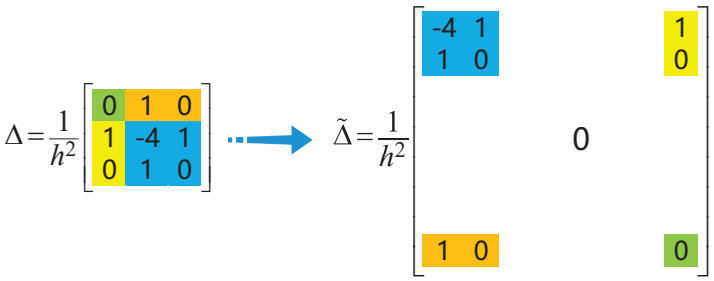

Figure 4: Expansion of the Laplacian convolution kernel to the image size: cyclically shift the original convolution kernel, so that the central element of the kernel is at (0, 0).

Note $\boldsymbol{p} = (p_1, p_2)$, $\boldsymbol{\lambda_1} = (\lambda_{11}, \lambda_{12})$. We calculate

$$B(i,j) = \alpha\big(u(i,j) - v(i,j)\big) - \gamma_1\big(\partial_1^- p_1^k(i,j) + \partial_2^- p_2^k(i,j)\big) - \big(\partial_1^- \lambda_{11}^{k-1}(i,j) + \partial_2^- \lambda_{12}^{k-1}(i,j)\big),$$

and we have noticed that the size of q is bigger than the Laplacian convolution kernels $\triangle$, which is usually the case in practice, the convolution kernels need to be expanded to the image size and padded according to Fig. 4. In fact, cyclic convolution with the expanded kernel is equivalent to cyclic convolution with initial convolution kernel.

For the Fourier-based convolution to satisfy Neumann boundary condition, the image needs to be expanded and padded as in [48]. By applying FFTs, one gets

$$(\lambda + \alpha)\mathcal{F}q^k(i,j) - \gamma_1 \mathcal{F}(\tilde{\triangle})\mathcal{F}q^k(i,j) = \mathcal{F}B(i,j), \tag{3.3}$$

where $\mathcal{F}(\cdot)$ is the two-dimensional Fourier transform function. Then once $\mathcal{F}q^k(i,j)$ is calculated, $q^k(i,j)$ could be obtained using the two-dimensional inverse Fourier transform.

After q and $\boldsymbol{p}$ are updated, the Lagrange multiplier $\boldsymbol{\lambda_1}$ will be advanced

$$\lambda_{11}^k(i,j) = \lambda_{11}^{k-1}(i,j) + \gamma_1\left(p_1^k(i,j) - \partial_1^+ q^k(i,j)\right), \tag{3.4}$$

$$\lambda_{12}^k(i,j) = \lambda_{12}^{k-1}(i,j) + \gamma_1\left(p_2^k(i,j) - \partial_2^+ q^k(i,j)\right). \tag{3.5}$$

This iterative method for approximating the saddle point of the functional (3.1) is given in Algorithm 3.1. We mark it as Algorithm ALM-FFTs.

Algorithm 3.1 ALM for the minimization of the proposed model (Eq. 2.2).

1: Initialization: $q^0 = (\alpha/(\lambda+\alpha))(u-v)$, $\boldsymbol{p}^0$ and ${\boldsymbol{\lambda_1}}^0$.
2: **for** $k \geq 1$ **do**
3: Compute the minimizer $\boldsymbol{p}^k$ by (3.2), q^k by (3.3), for the associated subproblems with the fixed Lagrangian multiplier ${\boldsymbol{\lambda_1}}^{k-1}$.
4: Update the Lagrange multiplier ${\boldsymbol{\lambda_1}}^k$ by (3.4) and (3.5).

5: Measure the relative error of the solution q^k (Eq. (4.1)) and stop the iteration if they are smaller than a given threshold ϵ_r. Mark the solution as q^*.
6: **end for**
7: Calculate $s = q^* + v$, and adjust the value range of s to $[0, 1]$.

The following convergence result for the sequence $\{q^k\}$ could be proved using a similar procedure as [3, 46, 52, 53].

Theorem 3.1. *Suppose* $(q^*, \boldsymbol{p}^*; {\boldsymbol{\lambda_1}}^*)$ *is a saddle-point of the augmented Lagrangian functional* $\mathcal{L}(q, \boldsymbol{p}; \boldsymbol{\lambda_1})$. *Let* $(q^k, \boldsymbol{p}^k; {\boldsymbol{\lambda_1}}^k), k = 1, 2, 3, \ldots$ *be the sequence generated by Algorithm* 3.1. *If* $\gamma_1 + 2\eta > 1/a$, *then one gets*

$$\lim_{k\to\infty} q^k = q^*.$$

For a non-convex optimization problem, only local optimal solutions can be obtained in most cases. In our convergence study, we prove that if a saddle-point of the augmented Lagrangian functional exists, then the sequence generated by the proposed algorithm converges to the saddle-point when $\gamma_1 + 2\eta > 1/a$. However, as our model is non-convex, the obtained minimizer might still be a local optimal solution.

4. Numerical experiments

In this section, we present numerical results obtained by applying our proposed model with ALM-FFTs to fuse images from the infrared and visible ones. Our objective is to fuse the visible image and the infrared image, while preserving the thermal radiation information from the infrared image and retaining the detailed appearance information from the visible image. Additionally, our model aims to eliminate noise, promote sharp edges and reduce the staircase effect. We compare the performance of our model with other image fusion techniques such as the TV model, the Huber model, LP, DWT, DTCWT, CVT, NSCT, MSVD, GFF. For all the numerical experiments, we use the following stopping criterion:

$$\frac{\|q^k - q^{k-1}\|_F}{\|q^{k-1}\|_F} < 1 \times 10^{-7}. \tag{4.1}$$

To monitor the convergence of the iterative process, we check the following relative residual as in [39]:

$$R^k = \frac{\|\boldsymbol{p}^k - \nabla q^k\|_F}{|\Omega|}, \tag{4.2}$$

the relative error of the Lagrange multipliers ${\boldsymbol{\lambda_1}}^k$

$$L^k = \frac{\|{\boldsymbol{\lambda_1}}^k - {\boldsymbol{\lambda_1}}^{k-1}\|_F}{\|{\boldsymbol{\lambda_1}}^{k-1}\|_F}, \tag{4.3}$$

and the relative error of the solution q^k

$$\frac{\|q^k - q^{k-1}\|_F}{\|q^{k-1}\|_F}, \tag{4.4}$$

where $\|\cdot\|_F$ is the Frobenius norm on Ω, $|\Omega|$ is the area of domain, and k refers to the iteration number. Moreover, for the purpose of presentation, all the above quantities are shown in log-scale in the figures. And we set the mesh size $h = 0.01$, the parameters $a = 2$, $\gamma_1 = 0.5$ in all experiments. The experiments are programmed in MatLab 2015a on an Intel(R) Core(TM) i7-8550U CPU@ 1.80GHz 2.00GHz desktop with 16.0 GB RAM.

When there is no standard reference image, eight quantitative metrics are used to evaluate the performance of different fusion methods. They include entropy (EN), standard deviation (SD), the mean structural similarity index measure (MSSIM), spatial frequency (SF), correlation coefficient (CC), average gradient (AG), peak signal-to-noise ratio (PSNR), and edge retentiveness ($Q^{AB/F}$) [17, 38, 55].

To be specific, we recall these metrics one by one in what follows. The entropy (EN) represents the average amount of information in an image, and it is defined as follows

$$EN = -\sum_{i=0}^{L-1} P_i \log_2 P_i,$$

where L is the total gray level of the image, P_i is the ratio of the number of pixels N_i with the gray value of i to the number of pixels N. The larger the value of EN, the richer the information of the fused image.

The standard deviation (SD) is an objective evaluation index to measure the richness of image information. A larger SD indicates the higher contrast of the fused image,

$$SD = \sqrt{\frac{M \times N}{\sum_{i=1}^{M}\sum_{j=1}^{N}(u(i,j) - \bar{u})^2}},$$

where u is the fused image, $\bar{u}$ is the mean of the signal u, and M, N denote the dimensions of the signal u.

The mean structural similarity index measure (MSSIM) is used to describe how the structure inside an image is maintained. Specifically, the larger the MSSIM, the better the structure is kept. This measure is defined as follows:

$$MSSIM = \frac{SSIM(A,F) + SSIM(B,F)}{2},$$

where

$$SSIM(A,F) = \frac{2\bar{F}\bar{A} + c_1}{\bar{F}^2 + \bar{A}^2 + c_1} \cdot \frac{2\sigma_F\sigma_A + c_2}{\sigma_F^2 + \sigma_A^2 + c_2} \cdot \frac{\sigma_{FA} + c_3}{\sigma_F\sigma_A + c_3},$$

and A represents the source image, F is the fused image, $\bar{F}$ is the mean of the signal F, σ_F is the standard deviation of the signal F, σ_{FA} is the covariance of the signal F and A, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$, $c_3 = c_2/2$, $k_1 = 0.01$, $k_2 = 0.03$, $L = 255$.

The spatial frequency (SF) is based on the image gradient and it reflects the image detail and texture sharpness level. The average gradient (AG) represents the ability to express the texture and detail of the fused image and can be used to evaluate the sharpness of the image. Both SF and AG can be used to measure the edge texture information, and the larger the values of SF or AG, the richer the edge texture information depicted in the fused image. They are given as follows:

$$SF = \sqrt{\frac{1}{M \times N}\left(\sum_{i=1}^{M}\sum_{j=2}^{N}\big(F(i,j) - F(i,j-1)\big)^2 + \sum_{i=2}^{M}\sum_{j=1}^{N}\big(F(i,j) - F(i-1,j)\big)^2\right)},$$

$$AG = \frac{1}{M \times N}\sum_{i=2}^{M}\sum_{j=2}^{N}\sqrt{\frac{(F(i,j) - F(i,j-1))^2 + (F(i,j) - F(i-1,j))^2}{2}},$$

where F is the fused image, and M, N denote the dimensions of the signal F.

The correlation coefficient (CC) measures the degree of linear correlation between a fused image and the source images, that is, the infrared and visible images. CC is defined as

$$CC = \frac{\rho_{AF} + \rho_{BF}}{2},$$

where ρ_{AF} is the correlation coefficient of the signal A and F. The larger CC indicates the fused image is more closely related to the source images.

Peak signal-to-noise ratio (PSNR) represents the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the representation fidelity of the signal. Since many signals have a wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. PSNR could reflect the distortion degree of the fused image and is given as

$$PSNR = 10\log_{10}\frac{2r^2}{MSE(A,F) + MSE(B,F)},$$

where r is the peak value of the fused image, and MSE represents the mean square error.

$Q^{AB/F}$ uses local metrics to measure the degree of the edges transferred from the source images to the fused image, and it is defined as follows:

$$Q^{AB/F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\big(Q^{AF}(i,j)\omega^A(i,j) + Q^{BF}(i,j)\omega^B(i,j)\big)}{\sum_{i=1}^{M}\sum_{j=1}^{N}\big(\omega^A(i,j) + \omega^B(i,j)\big)},$$

where $Q^{AF}(i,j) = Q_g^{AF}(i,j)Q_a^{AF}(i,j)$, $Q_g^{AF}(i,j)$ and $Q_a^{AF}(i,j)$ are the edge strength and orientation values at the location (i,j), $\omega^A(i,j)$ and $\omega^B(i,j)$ denote the weights of $Q^{AF}(i,j)$ and $Q^{BF}(i,j)$.

We first consider the influence of the parameters η on the performance of our model for the image "Plane" with $\lambda = 2 \times 10^3, \alpha = 2 \times 10^3$ in Fig. 5. As η increases, the fusion image becomes smoother. For instance, when $\eta = 0.5$, the regular term in our model

Table 2: The evaluation index values with different parameters by our model for the image in Fig. 5.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
$\eta = 0$	2.6625	15.5342	0.4658	6.9999	0.4983	7.2172	14.1989	0.4590
$\eta = 0.01$	2.6613	15.4808	0.4663	6.9687	0.4986	7.2025	14.2036	0.4599
$\eta = 0.1$	2.6553	15.1442	0.4685	6.9051	0.4994	7.1781	14.2389	0.4819
$\eta = 0.5$	2.6386	14.3236	0.4776	6.8185	0.4988	7.0884	14.8042	0.5407

becomes strictly convex, resulting in a heavily blurred tail boundary on the plane in the fusion image. Table 2 clearly illustrates that smoother images yield higher values for MSSIM, PSNR and $Q^{AB/F}$. This suggests that these three metrics favor over-smoothed results.

To show the convergence of the iterative process of the our algorithm with different η, we present the plots of the relative residual R^k (4.2), the relative error in the Lagrange multiplier L^k (4.3), the relative error of q^k (4.4), and the energy $E(q^k)$ versus iteration for the experiment "Plane" in Fig. 6. These plots demonstrate the convergence of the iterative process, which indicates that a saddle point of the augmented Lagrangian functional and thus a minimizer of the image fusion model is being approached.

In the following experiments, we fix $\eta = 1 \times 10^{-4}$.

To demonstrate how our model promotes sharp boundaries, we apply it to the image "M3FD" with $\lambda = 2 \times 10^3, \alpha = 2 \times 10^3$, as shown in Fig. 7. We compare our model with the TV model and the Huber model for their performance. All the three models generate good fused images, however, they do show difference in generating fine details. To show this, we present the zoomed-in parts of the obtained images. The appearance of pedestrians demonstrates that our model could produce clear edges while both the TV model and the Huber model yield fuzzy edges. This is due to the fact that the potential function of our regularizer presents a slower growth rate than those of the other two models, which weakens the competition of the regularization term with the fitting term, allows relatively large jumps, and thus helps preserve sharp edges.

In Fig. 8, we present another experiment "Lamplight" to show that our model can suppress the staircase effect while removing noise. From the zoomed-in images, one can observe that both our model and the Huber model could yield smooth patches for the vehicle while the TV model shows the block phenomenon. This justifies that our model is capable of reducing the staircase effect during the image fusing processing. As discussed before, this feature is caused by our specifically designed regularizer which presents as the Tikhonov regularizer for regions with relatively small gradients. In this experiment "Lamplight", the infrared image is corrupted by Gaussian noise ($\sigma = 0.01$), and we use $\lambda = 2 \times 10^2, \alpha = 3 \times 10^2$.

We present the comparison of the quantitative metrics values of those fused images, including EN, SD, MSSIM, SF, CC, AG, PSNR, and $Q^{AB/F}$ in Tables 3 and 4. It can be seen from the above experiments that the images processed by our model could al-

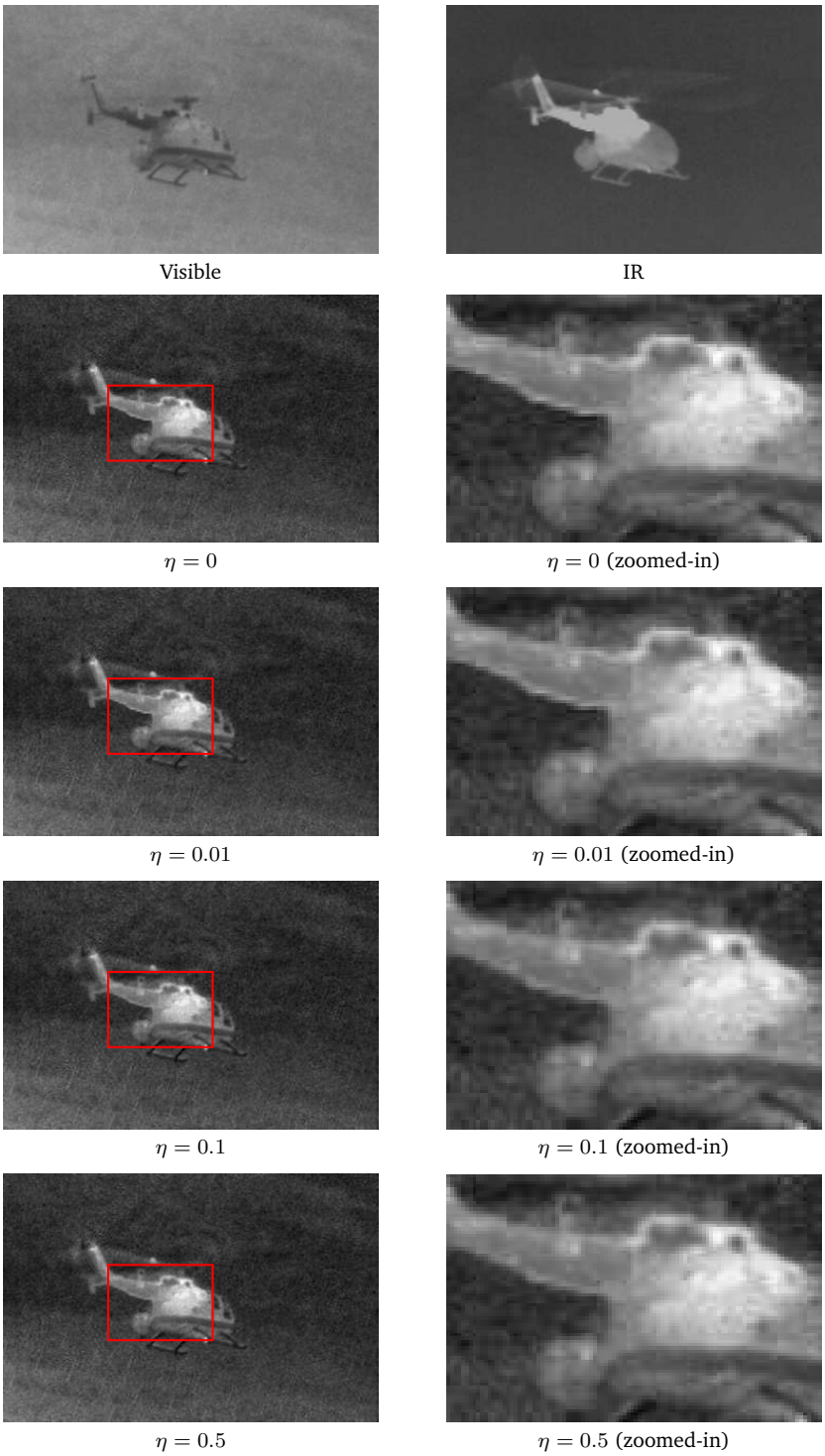

Figure 5: A visible image, an infrared image, the fused images and zoomed-in patches by our model with different η, respectively.

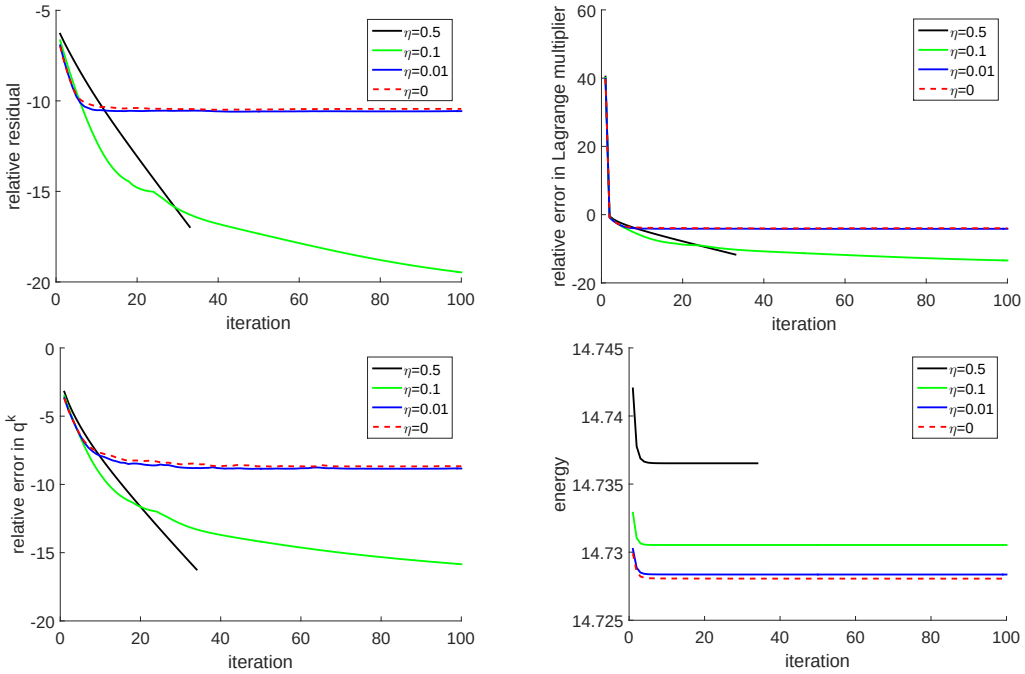

Figure 6: The plots of the relative residuals R^k (4.2), the relative errors in Lagrange multiplier L^k (4.3), the relative error in q^k (4.4), and the energy $E(q^k)$ versus iteration for the experiment "Plane" using our model with different algorithms. All the quantities are presented in $\log$-scale.

ways get a better visual effect, but the corresponding image quantitative metrics values may not be optimal. Since these quantitative metrics are based on the least squares measurement, they are more inclined to over-smoothed results.

In order to observe the influence of the parameters λ and α on the performance of our model, we consider an experiment "Kaptein" with different values of λ and α in Fig. 9. We mainly compare two cases. One case is to set $\lambda = 0$ while varying the value of α. Then our model imposes no constraint on the difference between the pixel intensities of s and v. As can be seen from the images on the left, when α is relatively small, such as $\alpha = 1$ and $\alpha = 10$, the fused image is close to the visible image; as α gradually increases, the fused image shows more and more information from the infrared image. Another case is to examine the effect of λ on the fused image by fixing $\alpha = 2\times10^3$. It can be seen from the images on the right column that when λ is relatively small, such as $\lambda = 2$, $\lambda = 20$ and $\lambda = 2 \times 10^2$, the fused image is close to the infrared image; as λ gradually increases, the fused image presents more and more information from the visible image. From Fig. 9 or Table 5, we can see that the image fusion works best when $\lambda = \alpha = 2 \times 10^3$. In fact, for most experiments, more pleasant fused images could be obtained by taking λ and α with the same order of magnitude.

In Figs. 10 and 11, we compare our method with other image fusion techniques (LP, DWT, DTCWT, CVT, NSCT, MSVD, GFF) on a set of image data-sets without standard reference images. The results show that our model can generate fused images


Visible

Visible (zoomed-in)


IR


IR (zoomed-in)


TV


TV (zoomed-in)


Huber


Huber (zoomed-in)


Ours


Ours (zoomed-in)

Figure 7: A visible image, an infrared image, the fused images and zoomed-in patches by the TV model, Huber model and our model, respectively.

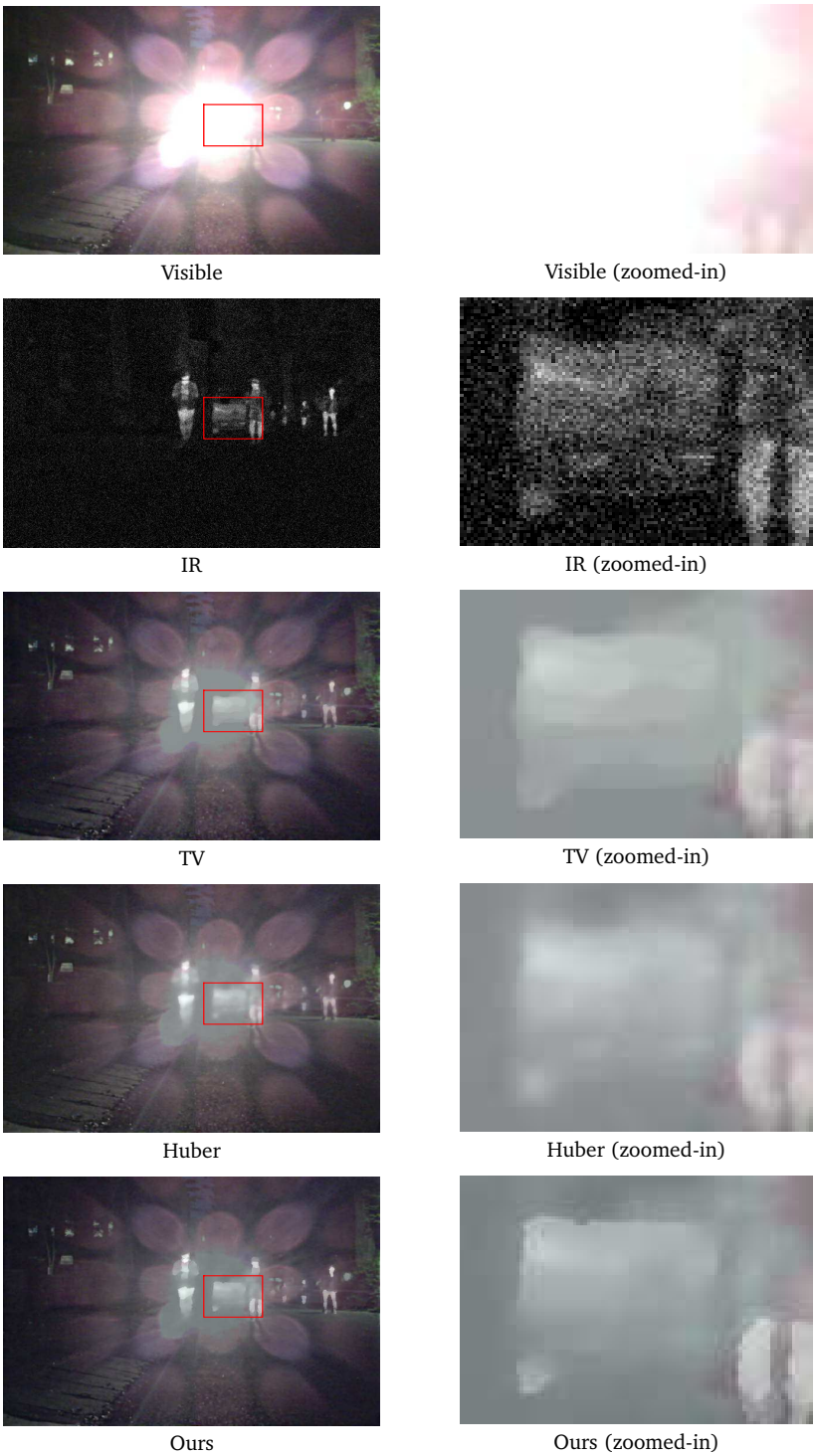

Figure 8: A visible image, an infrared image, the fused images and zoomed-in patches by the TV model, Huber model and our model, respectively.

Figure 9: A visible image, an infrared image and the fused images by our model and ALM-FFTs with different parameters λ and α.

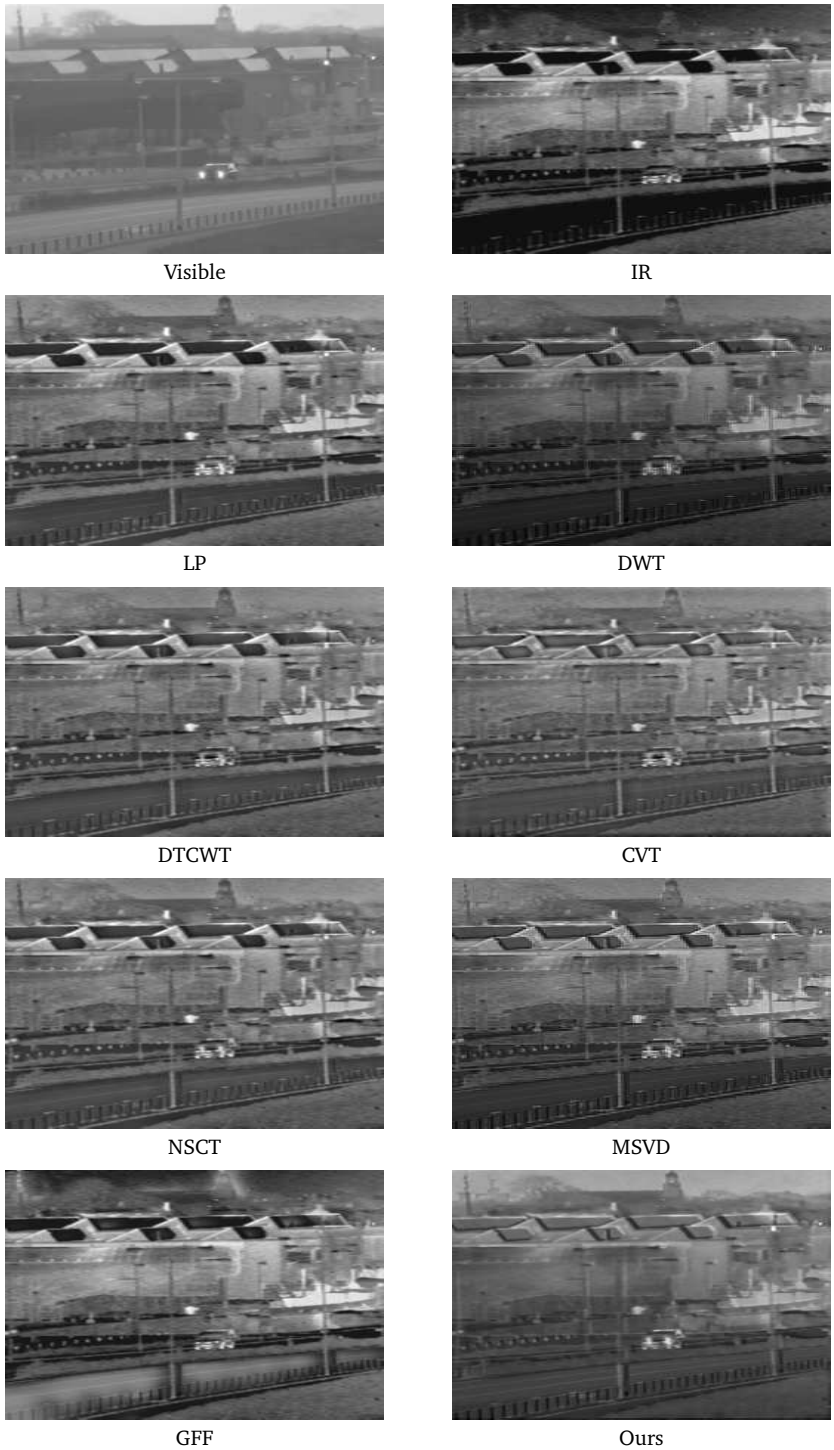

Figure 10: The visible image, the infrared image, the images fused using eight methods respectively (LP, DWT, DTCWT, CVT, NSCT, MSVD, GFF and our method).

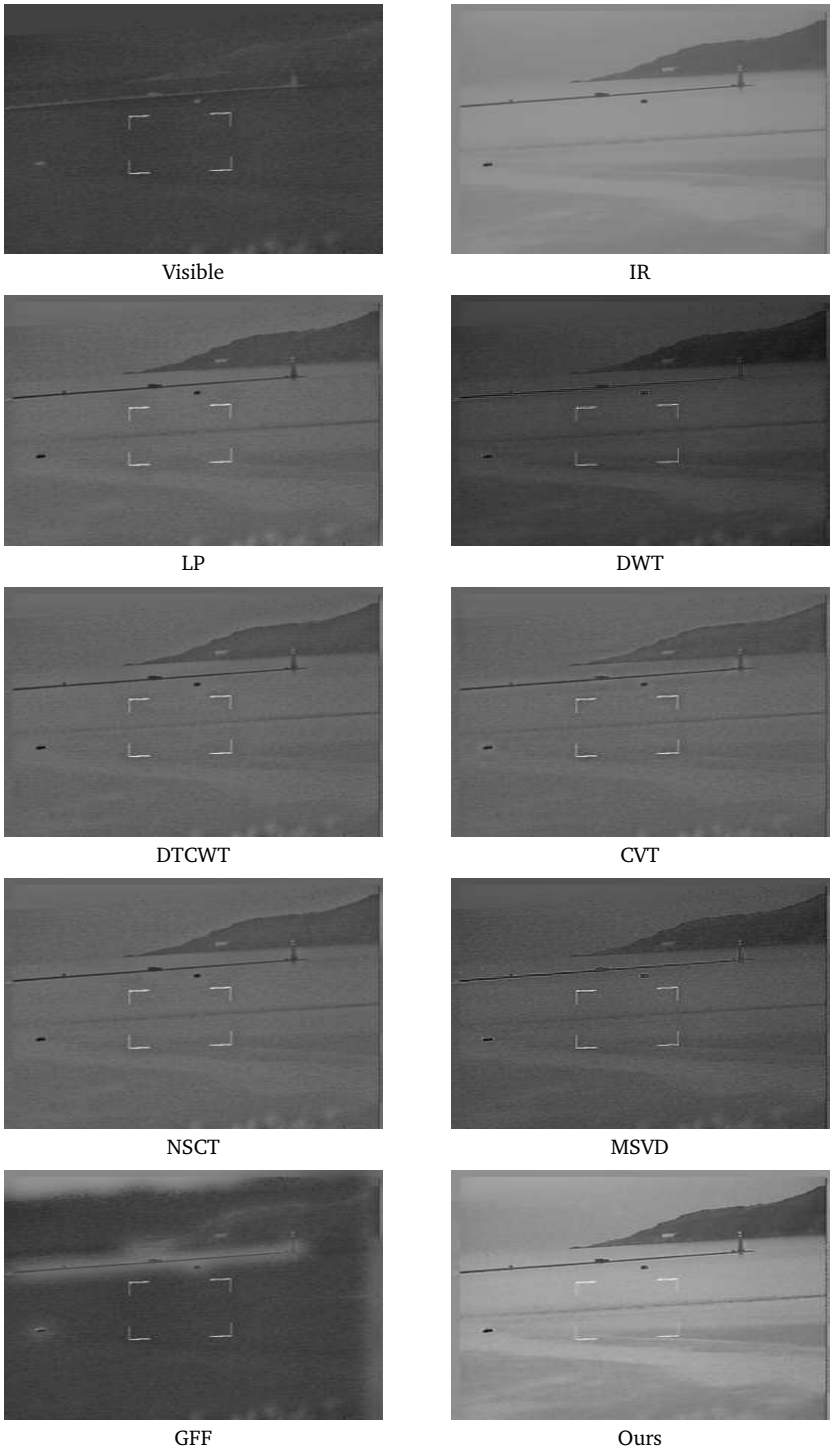

Figure 11: The visible image, the infrared image, the images fused using eight methods respectively (LP, DWT, DTCWT, CVT, NSCT, MSVD, GFF and our method).

Figure 12: The visible image, the infrared image, the images fused using eight methods respectively (IFCNN, SeAFusion, SDNet, RFN-Nest, DenseFuse, PMGI, FusionGAN and our method).

Table 3: The evaluation index values with different image fusion methods for the image in Fig. 7.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
TV	6.9199	31.2848	0.7015	6.2344	0.7340	2.1132	19.4544	0.3314
Huber	6.9183	31.2912	0.7013	6.1090	0.7344	2.0280	19.4124	0.3412
Ours	6.9724	32.5565	0.7004	5.9448	0.7348	2.0743	18.9231	0.3301

Table 4: The evaluation index values with different image fusion methods for the image in Fig. 8.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
TV	7.0862	37.5495	0.4563	5.0750	0.6536	2.4852	12.3558	0.7721
Huber	7.0084	35.8603	0.4553	4.6989	0.6629	2.3589	12.3597	0.7887
Ours	6.9369	34.1111	0.4415	5.2257	0.6660	2.2963	12.3880	0.7845

Table 5: The evaluation index values with different parameters by our model for the image in Fig. 9.

λ	α	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
0	1	6.8887	35.8104	0.4198	6.4637	0.4594	3.1561	14.9247	0.4070
0	10	6.7089	27.2247	0.3874	5.3021	0.4323	2.6202	14.0369	0.4536
0	100	6.4197	25.9407	0.3641	6.4163	0.3953	2.4459	12.1552	0.5046
0	1000	6.5508	29.4900	0.3832	6.2351	0.3897	2.5528	12.6723	0.5309
2	2000	6.6853	32.5063	0.3996	6.3454	0.3912	2.7112	13.2749	0.4997
20	2000	6.6678	32.3565	0.4046	6.3430	0.3975	2.7206	13.3274	0.4984
200	2000	6.6041	31.3398	0.4516	6.3719	0.4583	2.8342	13.7979	0.4844
2000	2000	6.8154	38.7321	0.5800	6.9210	0.6050	3.5985	16.1925	0.3510

Table 6: The evaluation index values with different image fusion methods for the image in Fig. 10.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
LP	7.2195	38.0366	0.4685	18.3916	0.4879	8.7048	15.6268	0.0875
DWT	7.1181	35.1982	0.4601	19.5384	0.4848	9.4372	16.1021	0.1194
DTCWT	7.0409	33.5963	0.4651	17.9801	0.4965	8.4826	15.9943	0.1009
CVT	6.9217	30.3511	0.4798	17.3409	0.5246	8.3842	16.0801	0.1257
NSCT	7.0791	34.2595	0.4710	18.1242	0.5001	8.5865	15.8084	0.0883
MSVD	7.1487	36.4634	0.5257	20.6421	0.5468	9.7427	15.0372	0.1429
GFF	7.3749	44.1471	0.3185	18.3067	0.3027	8.8080	13.2561	0.1056
Ours	7.0438	36.6781	0.5537	11.9666	0.5742	5.6163	15.0443	0.4127

with better visual effects than other techniques. To quantify the comparison of those image fusion techniques, we provide eight quantitative metrics values to conduct these techniques in Tables 6 and 7. These tables illustrate that our method produces higher quantitative metrics values than most of the existing techniques. In these experiments, we take $\lambda = 5 \times 10^3, \alpha = 3 \times 10^3$ in image "b07", $\lambda = 5 \times 10^3, \alpha = 2 \times 10^3$ in image "b08".

Table 7: The evaluation index values with different image fusion methods for the image in Fig. 11.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
LP	4.8541	9.9810	0.5992	8.3952	0.5223	3.7192	16.6710	0.0996
DWT	4.8661	9.4888	0.5934	8.5490	0.5105	3.8298	16.4292	0.1184
DTCWT	4.7671	8.9947	0.6018	8.3277	0.5202	3.6646	16.5586	0.1087
CVT	4.7397	8.4963	0.6113	8.2225	0.5335	3.6812	16.4083	0.1193
NSCT	4.7890	9.0892	0.6057	8.2878	0.5276	3.6520	16.6430	0.1006
MSVD	5.1250	10.9808	0.5937	11.0383	0.5555	5.0746	16.0594	0.1621
GFF	5.3948	18.3679	0.1923	7.7255	0.1497	3.7157	12.6442	0.2530
Ours	5.8148	20.9479	0.4691	10.0981	0.4467	3.4459	12.6918	0.3777

Table 8: The evaluation index values with different image fusion methods for the image in Fig. 12.

Methods	EN	SD	MSSIM	SF	CC	AG	PSNR	$Q^{AB/F}$
IFCNN	2.8638	22.5367	0.3456	7.6315	0.3617	6.1310	14.7549	0.1383
SeAFusion	2.9186	27.7988	0.3604	7.3859	0.3552	5.8535	13.5210	0.1953
SDNet	2.6417	12.7151	0.3649	6.7195	0.4069	5.3739	14.9704	0.1821
RFN-Nest	2.7424	15.1628	0.3118	3.3264	0.3950	2.7195	15.5064	0.4498
DenseFuse	2.7404	16.0193	0.3227	4.5318	0.3971	3.6258	15.4969	0.2749
PMGI	2.7938	16.1761	0.3867	4.5228	0.4357	3.6640	15.5522	0.3071
FusionGAN	2.6941	14.5424	0.3501	3.6536	0.3679	2.9907	12.0082	0.5503
Ours	2.6329	11.9441	0.3346	4.3093	0.4382	3.3466	15.6477	0.4316

In Fig. 12, we compare our method with a few popular deep learning-based methods, including IFCNN, SeAFusion, SDNet, RFN-Nest, DenseFuse, PMGI, and FusionGAN. The results show that our method can produce similar fusion results as SDNet, RFN-Nest, DenseFuse and PMGI. When compared to our model, image fusion using FusionGAN has a lower grayscale while image fusion using CNN and SeA Fusion has a stronger grayscale contrast. Moreover, it can be seen from the Table 8 that our method gives larger values of CC and PSNR than those deep learning-based method for the image in Fig. 12. In this experiment, we take $\lambda = 2 \times 10^3, \alpha = 3 \times 10^3$.

5. Conclusions

In this paper, we propose a non-convex first-order variational model for image fusion. Our model incorporates a novel regularizer that is able to eliminate noise, preserve sharp edges and reduce the staircase effect effectively. We prove that there exists a minimizer in $W^{1,2}(\Omega)$ for our model and also discuss its maximum-minimum principle. Augmented Lagrangian method (ALM) is used to design a fast algorithm to solve our model efficiently. We present numerical experiments to demonstrate the specific features of our model and compare it with other image fusion techniques.

Acknowledgments

This work was partially supported by the NSFC Projects No. 12001529, 12025104, 11871298, 81930119.

References

[1] T. ADAM AND R. PARAMESRAN, *Hybrid non-convex second-order total variation with applications to non-blind image deblurring*, Signal, Image Video P. 14 (2020), 115–123.
[2] G. BELLETTINI, V. CASELLES, AND M. NOVAGA, *The total variation flow in* $\mathbb{R}^n$, J. Differential Equations 184 (2002), 475–525.
[3] S. BOYD, N. PARIKH, AND E. CHU, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn. 3 (2011), 1–122.
[4] K. BREDIES, K. KUNISCH, AND T. POCK, *Total generalized variation*, SIAM J. Imaging Sci. 3 (2010), 492–526.
[5] P. J. BURT AND E. H. ADELSON, *The Laplacian pyramid as a compact image code*, IEEE Trans. Commun. 31 (1983), 532–540.
[6] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision 20 (2004), 89–97.
[7] T. CHAN, A. MARQUINA, AND P. MULET, *High-order total variation-based image restoration*, SIAM J. Sci. Comput. 22 (2000), 503–516.
[8] 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), 1964–1977.
[9] S. H. CHAN, R. KHOSHABEH, K. B. GIBSON, P. E. GILL, AND T. Q. NGUYEN, *An augmented Lagrangian method for total variation video restoration*, IEEE Trans. Image Process. 20 (2011), 3097–3111.
[10] C. CHEN, R. H. CHAN, S. MA, AND J. YANG, *Inertia proximal ADMM for linearly constrained separable convex optimization*, SIAM J. Imaging Sci. 8 (2015), 2239–2267.
[11] L. J. CHIPMAN, T. M. ORR, AND L. N. GRAHAM, *Wavelets and image fusion*, in: Proceedings of International Conference on Image Processing 3 (1995), 248–251.
[12] B. DACOROGNA, *Direct Methods in the Calculus of Variations*, Springer, 2008.
[13] A. L. DA CUNHA, J. ZHOU, AND M. N. DO, *Nonsubsampled contourilet transform: Filter design and applications in denoising*, in: Proceedings of IEEE International Conference on Image Processing 1 (2005), 1–749.
[14] A. L. DA CUNHA, J. ZHOU, AND M. N. DO, *The nonsubsampled contourlet transform: Theory, design, and applications*, IEEE Trans. Image Process. 15 (2006), 3089–3101.
[15] Y. FU AND X. WU, *A dual-branch network for infrared and visible image fusion*, in: Proceedings of 25th International Conference on Pattern Recognition (ICPR) (2021), 10675–10680.
[16] H. HARIHARAN, A. KOSCHAN, AND M. ABIDI, *The direct use of curvelets in multifocus fusion*, in: Proceedings of 16th IEEE International Conference on Image Processing (ICIP) (2009), 2185–2188.
[17] R. HOU, D. ZHOU, R. NIE, D. LIU, L. XIONG, Y. GUO, AND C. YU, *VIF-Net: An unsupervised framework for infrared and visible image fusion*, IEEE Trans. Comput. Imaging 6 (2020), 640–651.
[18] P. J. HUBER, *Robust estimation of a location parameter*, Ann. Math. Statist. 35 (1964), 73–101.

[19] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and Their Applications*, SIAM, 2000.
[20] N. KINGSBURY, *Image processing with complex wavelets*, Philos. Trans. Roy. Soc. A 357 (1999), 2543–2560.
[21] J. J. LEWIS, R. J. O'CALLAGHAN, S. G. NIKOLOV, D. R. BULL, AND N. CANAGARAJAH, *Pixel-and region-based image fusion with complex wavelets*, Inf. Fusion 8 (2007), 119–130.
[22] H. LI AND X. WU, *DenseFuse: A fusion approach to infrared and visible images*, IEEE Trans. Image Process. 28 (2019), 2614–2623.
[23] H. LI, X. WU, AND J. KITTLER, *RFN-Nest: An end-to-end residual fusion network for infrared and visible images*, Inf. Fusion 73 (2021), 72–86.
[24] S. LI, X. KANG, AND J. HU, *Image fusion with guided filtering*, IEEE Trans. Image Process. 22 (2013), 2864–2875.
[25] J. MA, C. CHEN, C. LI, AND J. HUANG, *Infrared and visible image fusion via gradient transfer and total variation minimization*, Inf. Fusion 31 (2016), 100–109.
[26] J. MA, W. YU, C. LI, P. LIANG, C. LI, AND J. JIANG, *FusionGAN: A generative adversarial network for infrared and visible image fusion*, Inf. Fusion 48 (2019), 11–26.
[27] Y. MA, J. CHEN, C. CHEN, F. FAN, AND J. MA, *Infrared and visible image fusion using total variation model*, Neurocomputing 202 (2016), 12–19.
[28] P. MARCELLINI, *A relation between existence of minima for non convex integrals and uniqueness for non strictly convex integrals of the calculus of variations*, in: Mathematical Theories of Optimization. Lecture Notes in Mathematics, J.P. Cecconi, T. Zolezzi (Eds.), Springer, 979 (1983), 216–231.
[29] A. MARQUINA AND S. J. OSHER, *Image super-resolution by TV-regularization and Bregman iteration*, J. Sci. Comput. 37 (2008), 367–382.
[30] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, AMS, 2001.
[31] C. B. MORREY, *Quasi-convexity and the lower semicontinuity of multiple integrals*, Pacific J. Math. 2 (1952), 25–53.
[32] V. NAIDU, *Image fusion technique using multi-resolution singular value decomposition*, Def. Sci. J. 61 (2011), 479–484.
[33] F. NENCINI, A. GARZELLI, S. BARONTI, AND L. ALPARONE, *Remote sensing image fusion using the curvelet transform*, Inf. Fusion 8 (2007), 143–156.
[34] M. NIKOLOVA AND M. K. NG, *Analysis of half-quadratic minimization methods for signal and image recovery*, SIAM J. Sci. Comput. 27 (2005), 937–966.
[35] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul. 4 (2005), 460–489.
[36] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D. 60 (1992), 259–268.
[37] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in Imaging*, Springer, 2009.
[38] C. SUN, C. ZHANG, AND N. XIONG, *Infrared and visible image fusion techniques based on deep learning: A review*, Electronics 9 (2020), 2162.
[39] X. C. TAI, J. HAHN, AND G. J. CHUNG, *A fast algorithm for Euler's elastica model using augmented Lagrangian method*, SIAM J. Imaging Sci. 4 (2011), 313–344.
[40] L. TANG, J. YUAN, AND J. MA, *Image fusion in the loop of high-level vision tasks: A semantic-aware real-time infrared and visible image fusion network*, Inf. Fusion 82 (2022), 28–42.

[41] A. TOET, *Image fusion by a ratio of low-pass pyramid*, Pattern Recognit Lett. 9 (1989), 245–253.
[42] L. VESE, *A study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optim. 44 (2001), 131–161.
[43] W. WANG AND F. CHANG, *A multi-focus image fusion method based on Laplacian pyramid*, J. Comput. 6 (2011), 2559–2566.
[44] Y. WANG, W. YIN, AND J. ZENG, *Global convergence of ADMM in nonconvex nonsmooth optimization*, J. Sci. Comput. 78 (2019), 29–63.
[45] J. WEICKERT, *Scale-space properties of nonlinear diffusion filtering with a diffusion tensor*, Laboratory of Technomathematics, University of Kaiserslautern, (1994), 1–29.
[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), 300–339.
[47] W. YANG, Z. HUANG, AND W. ZHU, *Image segmentation using the Cahn-Hilliard equation*, J. Sci. Comput. 79 (2019), 1057–1077.
[48] W. YANG, Z. HUANG, AND W. ZHU, *A first-order Rician denoising and deblurring model*, Inverse Probl. Imaging 17 (2023), 1139–1164.
[49] H. ZHANG AND J. MA, *SDNet: A versatile squeeze-and-decomposition network for real-time image fusion*, Int. J. Comput. Vis. 129 (2021), 2761–2785.
[50] H. ZHANG, H. XU, Y. XIAO, X. GUO, AND J. MA, *Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity*, in: Proceedings of the AAAI Conference on Artificial Intelligence 34 (2020), 12797–12804.
[51] Z. ZHAO, S. XU, J. ZHANG, C. LIANG, C. ZHANG, AND J. LIU, *Efficient and model-based infrared and visible image fusion via algorithm unrolling*, IEEE Trans. Circuits Syst. Video Technol. 32 (2021), 1186–1196.
[52] W. ZHU, *A first-order image denoising model for staircase reduction*, Adv. Comput. Math. 45 (2019), 3217–3239.
[53] W. ZHU, *A first-order image restoration model that promotes image contrast preservation*, J. Sci. Comput. 88 (2021), 1–23.
[54] W. ZHU, X. C. TAI, AND T. CHAN, *Augmented Lagrangian method for a mean curvature based image denoising model*, Inverse Probl. Imaging 7 (2013), 1409–1432.
[55] Z. ZHU, H. WEI, G. HU, Y. LI, G. QI, AND N. MAZUR, *A novel fast single image dehazing algorithm based on artificial multiexposure image fusion*, IEEE Trans. Instrum. Meas. 70 (2021), 1–23.