Image fusion and denoising using fractional-order gradient information

Mei, Jin-Jin; Dong, Yiqiu; Huang, Ting-Zhu

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Image fusion and denoising using fractional-order gradient information *

Jin-Jin Mei† Yiqiu Dong‡ Ting-Zhu Huang§

Abstract

Image fusion and denoising are significant in image processing because of the availability of multi-sensor and the presence of the noise. The first-order and second-order gradient information have been effectively applied to deal with fusing the noiseless source images. In this paper, due to the advantage of the fraction-order derivative, we first integrate the fractional-order gradients of noisy source images as the target fraction-order feature, and make it fit with the fractional-order gradient of the fused image. Then we introduce the total variation (TV) regularization for removing the noise. By adding the data fitting term between the fused image and a preprocessed image, a new convex variational model is proposed for fusing the noisy source images. Furthermore, an alternating direction method of multiplier (ADMM) is developed for solving the proposed variational model. Numerical experiments show that the proposed method outperforms the conventional total variation methods for simultaneously fusing and denoising.

Keywords: image fusion and denoising, alternating direction method of multiplier, inverse problem, fractional-order derivative, structure tensor

1 Introduction

Image fusion is an important issue due to the availability of multi-sensor data in various fields such as concealed weapon detection, remote sensing, medical diagnosis, defect inspection, and military surveillance [1–3]. Image fusion is to synthesize several source images of the same scene into a single image which contains much more visual information. But, the observed source images are inevitably corrupted by the noise during the process of image collection, acquisition, transmission and storage. For obtaining a high-quality image, it is necessary to simultaneously fuse and denoise the source images.

Researchers have paid great attention to image fusion and studied various efficient algorithms. According to the different degree of information extracted from the source images, image fusion methods are typically divided into three levels: pixel-level methods, feature-level methods and decision-making level methods. The pixel-level methods are to directly estimate the pixel values of the fused image by integrating the pixel values of several source images with a feature selection rule [4,5]. In this paper, since it preserves more original information from the source images [6,7],

*This research was supported by 973 Program (2013CB329404), NSFC (61370147, 61402082, 11401081) and the Fundamental Research Funds for the Central Universities (ZYGX2016J129).

†School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China (meijinjin666@126.com).

‡Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark (yido@dtu.dk). The work was supported by Advanced Grant 291405 from the European Research Council.

§School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China (tingzhuhuang@126.com).

we devote to the study of the image fusion methods based on the pixel-level. Generally, the pixel-level fusion approaches mainly consist of spatial domain fusion methods [8–12] and transform domain fusion methods [13–19]. Among the spatial domain fusion methods, some variational methods combined the structure tensor with gradient information have been studied for dealing with the image fusion problem [11, 12, 20].

Before introducing the pixel-level fusion approaches, we firstly review some notations with respect to the structure tensor. Suppose that $\Omega \subset \mathbb{R}^2$ is a connected bounded image domain with the compacted Lipschitz boundary, $u : \Omega \to \mathbb{R}$ is the unknown fused image, $f_i : \Omega \to \mathbb{R}$, $i = 1, 2, \cdots, l$ are the source images. According to [11, 21, 22], the structure tensor is used for depicting the local structure feature, which is defined as follows

$$G(x, y) \triangleq \begin{pmatrix} \sum_{i=1}^{l} \left(\frac{\partial f_i}{\partial x}\right)^2 & \sum_{i=1}^{l} \frac{\partial f_i}{\partial x}\frac{\partial f_i}{\partial y} \\ \sum_{i=1}^{l} \frac{\partial f_i}{\partial y}\frac{\partial f_i}{\partial x} & \sum_{i=1}^{l} \left(\frac{\partial f_i}{\partial y}\right)^2 \end{pmatrix}, \quad \forall (x, y) \in \Omega.$$

Then, the target gradient $V : \Omega \to \mathbb{R}^2$ consists of eigenvalues and eigenvectors of the structure tensor G. Since G is positive semi-definite, the target gradient V is formulated as

$$V \triangleq \sqrt{\lambda}\mathbf{e}$$

where λ denotes the largest eigenvalues of G which conveys the sharp information, $\mathbf{e}$ is the corresponding unit eigenvector which indicates the orientation maximizing the pixel-value fluctuations and satisfies that $\langle \mathbf{e}, \sum_{i=1}^{l} \nabla f_i \rangle \geq 0$. Based on the above target gradient of source images, a classic variational model is proposed as

$$\min_u \int_\Omega |\nabla u - V|^2 dx. \tag{1}$$

Note that the model (1) results in the decrease of the contrast and the solution is not unique. In [11], Piella combined the image enhancement techniques and proposed a variational model for integrating the salient structure of source images and enhancing the contrast of the fused image. The author derived the corresponding Euler-Lagrange equation and applied the gradient descent method for obtaining the unknown fused image. But due to the slow convergent speed, the gradient descent method is extremely time-consuming. In [23], the authors considered the first-order and second-order gradient information at different directions. By adding a fidelity term which requires that the fused image is close to a predefined image u_0. They presented a new variational approach (called FSGF for short),

$$\min_u \int_\Omega \left(|\nabla u - sv| + \alpha|\nabla^2 u - sw|\right) dx + \frac{\beta}{2} \int_\Omega |u - u_0|^2 dx \tag{2}$$

where $\alpha \geq 0, \beta > 0$, $s \geq 1$, v and w denote the target gradients which integrated the first-order and second-order gradient information of the source images, respectively. A new feature selection was used to construct the target gradient v and w. Note that if $s = 1$, the model (2) is just for image fusion; if $s > 1$, the model is for simultaneous image fusion and enhancement. Moreover, the split Bregman method [24–26] was applied to solve the model (2).

As far as we know, a few researchers have paid attention to study the effective variational approaches for fusing the noisy source images. In [20], combining with the TV regularization, Wang *et al*. proposed a variational model for fusing and denoising the noisy multi-focus images (just referred to as WFTV),

$$\min_u \int_\Omega \left(\alpha_1(x)(u - f_1)^2 + \alpha_2(x)(u - f_2)^2\right) dx + 2\beta \int_\Omega |\nabla u| dx \tag{3}$$

where $\alpha_1(x)$ and $\alpha_2(x)$ are two nonnegative weight functions such that $\alpha_1(x) + \alpha_2(x) = 1$, f_1 and f_2 are the noisy source images, β is the positive parameter which trades off the fidelity term for fusing the noisy multi-focus images and the TV regularization for smoothing the fused image. They applied the gradient descent method to deal with (3). In [27], the authors proposed a adaptive sparse representation model for image fusion and denoising. Assume that the high-quality image patches were classified into several categories in advance. Then, based on the gradient information of images, they learned a set of compacted sub-dictionaries from numerous high-quality image patches and designed an algorithm to select adaptively one of the sub-dictionaries. Their method was based on dictionary learning, not belong to the variational method. And it was fairly complicated and time consuming. However, in the following, we only focus on the variational methods for image fusion and denoising.

Recently, in [28], the authors illustrated the theoretical properties of the total fraction-order variation. Based on the the total fraction-order variation, their proposed model could preserve the edges and smoothness. Inspired by this advantage, we combine the fraction-order derivative and propose a novel variational model for image fusion and denoising. According to [29–31], we build the target fractional-order gradient information from the noisy source images by a simple salience criterion [23]. For fusing the source images, we require that the fractional-order gradient information of the fused image should be matched with the target fractional-order gradient information in the sense of L^2 norm. Meanwhile, the fused image is closed to the predefined image. Since the source images are corrupted by the noise, we add the TV regularization to the variational model for the regularity constrain. Furthermore, we prove the existence and uniqueness of the solution. Under the framework of ADMM [32–34], a fast algorithm is developed for solving the proposed model.

The rest of this letter is organized as follows. In the next section, we introduce the elementary properties of the fractional-order derivative and the framework of ADMM. In Section 3, we obtain the target the fractional-order feature from the noisy source images and construct a new variational model with the TV regularization for image fusion and denoising. Meanwhile we give the existence and the uniqueness of the solution. In Section 4, we apply the ADMM algorithm to solve the proposed model. Numerical experiments are shown in Section 5. Finally, we conclude the paper with a summary.

2 Preliminaries

In this section, for the completeness of the paper, we briefly review the fraction-order derivative and the classic ADMM algorithm.

2.1 Fractional-order derivative

As a generalization of the integer-order derivatives, the fractional-order derivative has been widely used in image processing [28, 35–37]. Firstly, the definition and property of the fractional-order derivative have been studied over the one-dimensional domain, see [29–31] and references therein. Assume that a fraction $\gamma \in \mathbb{R}^+$ and a function $h \in L^1([a,b];\mathbb{R})$, the left and right Riemann-Liouville (RL) fractional-order derivatives are respectively defined by

$$
\begin{aligned}
{}_aD_x^{\gamma}[h](x) &= \frac{1}{\Gamma(1-\gamma)}\frac{d}{dx}\int_a^x \frac{h(\tau)}{(x-\tau)^{\gamma}}d\tau, \quad x \in (a,b] \\
{}_xD_b^{\gamma}[h](x) &= \frac{-1}{\Gamma(1-\gamma)}\frac{d}{dx}\int_x^b \frac{h(\tau)}{(\tau-x)^{\gamma}}d\tau, \quad x \in [a,b)
\end{aligned}
$$

where $\Gamma(\gamma)$ represents a Gamma function. Based on the left and the right Riemann-Liouville fractional-order derivatives, the Riesz-RL (central) fractional-order derivative is given by

$$ {}_aD_b^{\gamma}[h](x) = \frac{1}{2}\left({}_aD_x^{\gamma}[h](x) - {}_xD_b^{\gamma}[h](x)\right), \quad x \in (a,b). $$

Similarly, the left, the right and the Riesz-Caputo fraction-order derivatives are defined by as follows

$$ \begin{aligned} {}_a^C D_x^{\gamma}[h](x) &= \frac{1}{\Gamma(1-\gamma)}\int_a^x \frac{h'(\tau)d\tau}{(x-\tau)^{\gamma}}, \quad x \in (a,b] \\ {}_x^C D_b^{\gamma}[h](x) &= -\frac{1}{\Gamma(1-\gamma)}\int_x^b \frac{h'(\tau)d\tau}{(\tau-x)^{\gamma}}, \quad x \in [a,b) \\ {}_a^C D_b^{\gamma}[h](x) &= \frac{1}{2}\left({}_a^C D_x^{\gamma}[h](x) - {}_x^C D_b^{\gamma}[h](x)\right), \quad x \in (a,b) \end{aligned} $$

where $h'(x)$ denotes the first-order derivative of the function h with respect to x. In what follows, we consider only the Riesz-RL and the Riesz-Caputo fractional-order derivatives. Note that when h is continuously differentiable and its first-order derivative is integrable, the RL and the Caputo fractional-order derivatives are equivalent under the homogeneous boundary condition [29, 38]. Due to the equivalency, both fractional-order derivatives are uniformly denoted by D^{γ}. Moreover, since there is a singularity at the endpoint, we set $h(a) = h(b) = 0$ for the fractional-order derivative.

In the discrete case, assume that $u \in \mathbb{R}^{n\times n}$ under Dirichlet boundary condition. According to [39,40], we give the discrete fractional-order derivatives at the point (i, j) with the order γ along the horizontal and the vertical direction, respectively

$$ \begin{aligned} D_h^{\gamma}u_{i,j} &= \frac{1}{2}\left(\sum_{k=0}^{i+1} c_k^{\gamma}u_{i-k+1,j} + \sum_{k=0}^{n-i+2} c_k^{\gamma}u_{i+k-1,j}\right), \quad i,j = 1,\cdots,n \\ D_v^{\gamma}u_{i,j} &= \frac{1}{2}\left(\sum_{k=0}^{j+1} c_k^{\gamma}u_{i,j-k+1} + \sum_{k=0}^{n-j+2} c_k^{\gamma}u_{i,j+k-1}\right), \quad i,j = 1,\cdots,n \end{aligned} $$

where $c_k^{\gamma} = (-1)^k\binom{\gamma}{k}$ for $k = 0, 1, \cdots, n+1$. Actually, we compute the coefficients c_k^{γ} by the following recursive relation

$$ c_0^{\gamma} = 1, \quad c_k^{\gamma} = \left(1 - \frac{1+\gamma}{k}\right)c_{k-1}^{\gamma}, \quad k = 1, 2, \cdots. $$

Let $c = c_0^{\gamma} + c_2^{\gamma}$, the discrete fractional-order derivative of u is formulated as

$$ D^{\gamma} : \mathbb{R}^{n\times n} \to \mathbb{R}^{2n\times n}, \quad D^{\gamma}u := \begin{pmatrix} Mu \\ uM \end{pmatrix} $$

where

$$ M = \begin{pmatrix} 2c_1^{\gamma} & c & c_3^{\gamma} & \cdots & c_n^{\gamma} \\ c & 2c_1^{\gamma} & \ddots & \ddots & \vdots \\ c_3^{\gamma} & \ddots & \ddots & \ddots & c_3^{\gamma} \\ \vdots & \ddots & \ddots & 2c_1^{\gamma} & c \\ c_n^{\gamma} & \cdots & c_3^{\gamma} & c & 2c_1^{\gamma} \end{pmatrix} \in \mathbb{R}^{n\times n}. $$

Due to the symmetry of the matrix M and the definition of the inner production on the matrix, i.e., $\langle A, B\rangle = tr(A^\top B) = (vecA)^\top vecB$ for all $A, B \in \mathbb{R}^{n\times n}$ where $vecA$ and $vecB$ are vectors by stacking all the columns of the matrix A and B respectively, we have the following equations

$$\begin{aligned}\langle Mu, v^1\rangle &= tr(u^\top(Mv^1)) = \langle u, Mv^1\rangle \\ \langle uM, v^2\rangle &= (vec(uM^\top))^\top vecv^2 = (vecu)^\top vec(v^2M) = \langle u, v^2M\rangle,\end{aligned}$$

for any $v^1, v^2 \in \mathbb{R}^{n\times n}$. Finally, we obtain that

$$\langle D^\gamma u, v\rangle = \langle Mu, v^1\rangle + \langle uM, v^2\rangle = \langle u, Mv^1 + v^2M\rangle, \text{ for all } v = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \in \mathbb{R}^{2n\times n}.$$

Here we define the transform of the discrete fraction-order operator as $(D^\gamma)^\top v = Mv^1 + v^2M$.

2.2 Classic ADMM

According to [32–34, 41], the ADMM technique is widely applied for dealing with the following general constrained optimization problem

$$\begin{aligned}&\min_{x\in X, y\in Y} f(x) + g(y) \\ &\text{s.t. } Ax + By = b\end{aligned} \tag{4}$$

where $f(x)$ and $g(y)$ are closed convex and lower semi-continuous functions, $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ are closed convex set, $A \in \mathbb{R}^{l\times m}$ and $B \in \mathbb{R}^{l\times n}$ are linear operators. By introducing a multiplier $p \in \mathbb{R}^l$ with the linear constraint $Ax + By = b$, the corresponding augmented Lagrangian function is given by

$$\mathcal{L}(x, y; \lambda) = f(x) + g(y) + p^\top(Ax + By - b) + \frac{\delta}{2}\|Ax + By - b\|_2^2 \tag{5}$$

where β is a positive penalty parameter. Since the x-subproblem and the y-subproblem are decoupled, the solution (x^{k+1}, y^{k+1}) of (4) is iteratively obtained by

$$\begin{cases} x^{k+1} = \arg\min_x f(x) + \frac{\delta}{2}\left\|Ax + By^k - b + \frac{p^k}{\delta}\right\|_2^2, \\ y^{k+1} = \arg\min_y g(y) + \frac{\delta}{2}\left\|Ax^{k+1} + By - b + \frac{p^k}{\delta}\right\|_2^2, \\ p^{k+1} = p^k + \tau\delta(Ax^{k+1} + By^{k+1} - b) \end{cases} \tag{6}$$

where $\tau > 0$ controls the convergence speed. According to [42], if $\tau \in \left(0, \frac{\sqrt{5}+1}{2}\right)$, the classic ADMM algorithm is convergent.

3 The Proposed Models

In this paper, we build the target fraction-order gradient feature by integrating several noisy source images. For obtaining the fused image, we assume that the fractional-order gradient information of the fused image matches with the target fractional-order gradient feature, and the fused image is closed to the predefined image u_0. Moreover, the TV regularization preserves the image edges

while removes the noise [43]. Inspired by the works [23, 29–31], by combining with the TV regularization, we propose a new variational model for simultaneously fusing and denoising (called FTVL1 for short),

$$\min_u \int_\Omega |\nabla u|\,dx + \alpha \int_\Omega |\nabla^\gamma u - w|\,dx + \frac{\beta}{2}\int_\Omega |u-u_0|^2\,dx, \tag{7}$$

where α, β are positive regularization parameters which control the balance between image fusion and denoising, w denotes the target fractional-order feature, u_0 represents the preprocessed image. Similarly, we also present a variational model by using the difference between of $\nabla^\gamma u$ and w in the sense of the L^2 norm (just referred to as FTVL2)

$$\min_u E(u) = \int_\Omega |\nabla u|\,dx + \frac{\alpha}{2}\int_\Omega |\nabla^\gamma u - w|^2\,dx + \frac{\beta}{2}\int_\Omega |u-u_0|^2\,dx, \tag{8}$$

Here, the first term $\int_\Omega |\nabla u|\,dx$ in the models (7) and (8) represents the total variation (TV) regularization, defined as follows

$$\int_\Omega |\nabla u|\,dx := \sup\left\{ \int_\Omega u\ \mathrm{div}\varphi\,dx \,\middle|\, \varphi \in (C_0^\infty(\Omega))^2, \|\varphi\|_\infty \le 1 \right\}$$

where $(C_0^\infty(\Omega))^2$ is the space of vector-valued functions with the compacted support in Ω. It is used for removing the Gaussian white noise in the source images.

The second terms $\int_\Omega |\nabla^\gamma u - w|\,dx$ in (7) and $\int_\Omega |\nabla^\gamma u - w|^2\,dx$ in (8) are used to integrate all the important salient features from the noisy source images. By combining the selection criterion [23], we construct the target fractional-order feature w from the noisy source images. Specifically, assume that two source images f_1 and f_2, we compute the fractional-order gradient $\nabla^\gamma f_i = \begin{pmatrix} Mf_i \\ f_i M \end{pmatrix}$, for $i = 1, 2$. According to the selection criterion [44, 45], we propose a relatively simple method for selecting the salient features from the noisy source images. The target fractional-order feature w is formulated as follows,

$$w = s\nabla^\gamma f_1 + (1-s)\nabla^\gamma f_2$$

where s denotes a binary mask for fusing the salient features from the noisy source images. By smoothing and thresholding, the binary mask is given as

$$s(x) = \begin{cases} 1, & \text{if } K * \tilde{s} > 0.5 \\ 0, & \text{otherwise,} \end{cases} \quad \text{with} \quad \tilde{s}(x) = \begin{cases} 1, & \text{if } K * (\nabla^\gamma f_1)^2 > K * (\nabla^\gamma f_2)^2 \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

where $*$ represents the convolution and K is a mean kernel function, i.e., if $y \in \Omega_x$, then $K(x, y) = 1/|\Omega_x|$; otherwise $K(x, y) = 0$. The domain Ω_x is a bounded neighborhood of x with area $|\Omega_x|$. Note that when the number of the source images is more than two, we utilize the recursive method and only incorporate two source images at one time. Furthermore, we choose the L^2 norm to measure the gap $\nabla^\gamma u - w$ in (8). This is because the L^2 norm comes from Gaussian distribution by taking $-\log$ operaton. Specially, we consider the image "Cameraman" as the test image. The noisy source images are obtained by blurred the left part and right part with the out-of-focus kernel, then added the Gaussian white noise with mean zero and standard variance 20, 30, respectively. The desired fused image u is equal to the truth image, the target fractional-order feature w is obtained by the above feature selection criterion. Figure 1 shows that the gap $\nabla^\gamma u - w$ in the horizontal direction for the Gaussian white noise with standard variance 20 and 30. Their corresponding histograms are shown in Figure 1(b) and (d) together with a fitting array containing random values

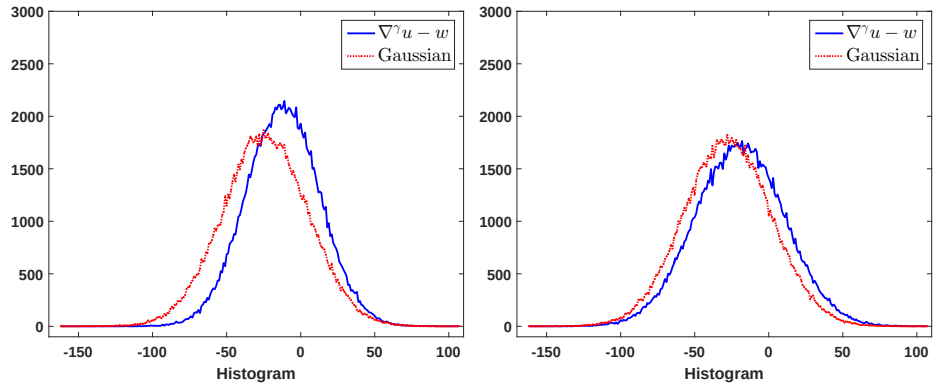


Figure 1. The corresponding histograms of the gaps $\nabla^\gamma u - w$ for the Gaussian white noise with standard variance 20 (in the left side) and 30 (in the right side), respectively.

follows Gaussian distribution. We find that the histogram of $\nabla^\gamma u - w$ is matched with the histogram of an array obeyed Gaussian distribution. Therefore, we prefer to apply the L^2 norm to measure the gap $\nabla^\gamma u - w$.

Furthermore, in order to obtain a better fused image, we assume that the fused image should be close to a preprocessed image u_0. Thus, a data-fitting term $\int_\Omega |u - u_0|^2\, dx$ is added to the proposed variational model. Since the preprocessed image u_0 contains all the information from the noisy source images, the proposed model improves effectively the quality of the fused images. Furthermore, for preserving the image contrast, we combine the average contrast (AC) [23] with the selection criterion. Specifically, if $AC(f_1) - AC(f_2) > 0.5$, then $u_0 = sf_1 + (1-s)f_2$ where s satisfies the equation (9); otherwise, u_0 is equal to the average of the noisy source images.

In the following, we give a theorem to illustrate the existence and uniqueness of the solution of the proposed model (8). First, according to [28], we give some notations and definitions of the space of functions with γ-bounded variation on Ω. Let $C_0^\ell(\Omega, \mathbb{R}^2)$ denote the space of ℓ-order continuously differentiable functions with compact support in $\Omega \subset \mathbb{R}^2$, $\mathcal{K}$ is the space of special test functions, which is formulated as

$$\mathcal{K} := \left\{ \varphi \in C_0^\ell(\Omega, \mathbb{R}^2) \,\middle|\, |\varphi| \leqslant 1 \text{ for all } x \in \Omega \right\}$$

where $|\varphi| = \sqrt{\varphi_1^2 + \varphi_2^2}$. Then, the fractional-order total variation of u is defined by

$$\int_\Omega |\nabla^\gamma u|\, dx := \sup \left\{ \int_\Omega -u\, div^\gamma \varphi\, dx \,\middle|\, \varphi \in \mathcal{K} \right\}$$

where $div^\gamma \varphi = \frac{\partial^\gamma \varphi_1}{\partial x} + \frac{\partial^\gamma \varphi_2}{\partial y}$, $\frac{\partial^\gamma \varphi_1}{\partial x}$ and $\frac{\partial^\gamma \varphi_2}{\partial y}$ denote the fractional-order derivatives along the x-direction and y-direction, respectively. Moreover, the space of functions with γ-bounded variation on Ω is given as follows,

$$BV^\gamma(\Omega) := \left\{ u \in L^1(\Omega) \,\middle|\, \int_\Omega |\nabla^\gamma u|\, dx < +\infty \right\}.$$

With the γ-BV norm $\|u\|_{BV^\gamma} = \|u\|_{L^1} + \int_\Omega |\nabla^\gamma u|\, dx$, the space $BV^\gamma(\Omega)$ is a Banach space.

Theorem 3.1. *Assume that $u_0 \in BV^\gamma(\Omega) \cap L^2(\Omega)$ and w has a finite vector-valued Radon measure, then the minimization problem* (8) *exists an unique minimizer $u^* \in BV^\gamma(\Omega) \cap L^2(\Omega)$.*

Proof. In order to prove the existence and uniqueness of the solution, we first let $\{u^k\} \subset BV^\gamma(\Omega) \cap L^2(\Omega)$ be a minimizing sequence of (8). Since $E(0)$ is finite, the minimum of the energy function $E(u)$ is also finite. In other words, there exists a positive constant C such that

$$E(u^k) = \int_\Omega |\nabla u^k| dx + \frac{\alpha}{2} \int_\Omega |\nabla^\gamma u^k - w|^2 dx + \frac{\beta}{2} \int_\Omega |u^k - u_0|^2 dx \leqslant C.$$

Since w has the finite vector-valued Radon measure, we have that

$$\int_\Omega |w| := \sup \left\{ \int_\Omega w \cdot \varphi \, dx \,\middle|\, \varphi \in \mathcal{K} \right\} < +\infty.$$

Due to the fact that $\int_\Omega |\nabla^\gamma u^k - w| \, dx := \sup \left\{ \int_\Omega \left(-u^k \, div^\gamma \varphi + w \cdot \varphi \right) dx \,\middle|\, \varphi \in \mathcal{K} \right\}$, then we obtain the following inequalities as follows,

$$\int_\Omega |\nabla^\gamma u^k| \, dx \leqslant \int_\Omega |\nabla^\gamma u^k - w| \, dx + \int_\Omega |w| \leqslant C \tag{10}$$

Since $\int_\Omega |\nabla u^k| \, dx := \sup \left\{ \int_\Omega u^k \, \mathrm{div} \varphi \, dx \,\middle|\, \varphi \in (C_0^\infty(\Omega))^2, \|\varphi\|_\infty \leq 1 \right\}$, we have

$$\int_\Omega |u^k - u_0|^2 \, dx \leqslant C. \tag{11}$$

According to the triangle inequality and $L^2(\Omega) \subset L^1(\Omega)$, we deduce that

$$\int_\Omega |u^k|^2 \, dx \leqslant C, \qquad \int_\Omega |u^k| \, dx \leqslant C. \tag{12}$$

Combining with (10), (11) and (12), we have that $\{u^k\}$ is bounded in $BV^\gamma(\Omega) \cap L^2(\Omega)$. Based on the weak* topology of $BV\gamma(\Omega)$ [28] and the reflexivity of $L^2(\Omega)$, there exists a subsequence converged to $u^* \in BV^\gamma(\Omega) \cap L^2(\Omega)$ such that

$$u^k \xrightarrow{BV^\gamma - w^*} u \quad \text{and} \quad u^k \xrightarrow{L^2(\Omega)} u.$$

According to the lower semi-continuity of bounded variation, γ-bounded variation and L^2-norm, we get

$$\begin{aligned}
&\int_\Omega |\nabla u^*| \, dx \leqslant \liminf_{k \to \infty} \int_\Omega |\nabla u^k| \, dx, \\
&\int_\Omega |\nabla^\gamma u^* - w|^2 \, dx \leqslant \liminf_{k \to \infty} \int_\Omega |\nabla^\gamma u^k - w|^2 \, dx, \\
&\int_\Omega |u^* - u_0|^2 \, dx \leqslant \liminf_{k \to \infty} \int_\Omega |u^k - u_0|^2 \, dx.
\end{aligned}$$

Therefore, by Fatou's lemma, we can conclude that $E(u^*) \leqslant \liminf_{k \to \infty} E(u^k)$ and u^* is a minimizer of (8). Moreover, since the energy function $E(u)$ is obviously strictly convex, the minimizer u^* is unique. □

4 Proposed Algorithms

In this section, we develop an effective numerical algorithm for solving the proposed model (8) based on the classic ADMM algorithm. In fact, some other numerical methods can also be used for addressing the proposed model, such as the gradient descent method [43], the primal-dual method [46, 47], the split Bregman method [25] and the alternating minimization method [48, 49]. Hereafter, we transform the proposed model into a discrete minimization. For the sake of simplicity, we still use the same notations in the model (8). The discrete minimization is rewritten as follows,

$$\min_{u\in\mathbb{R}^{n\times n}} \|Du\|_1 + \frac{\alpha}{2}\|D^\gamma u - w\|_F^2 + \frac{\beta}{2}\|u - u_0\|_F^2. \tag{13}$$

where u denotes the $n \times n$ gray-scale image and $\|\cdot\|_F$ represents the Frobenius norm. The discrete TV regularization term $\|Du\|_1$ under the zero boundary condition is formulated by

$$D : \mathbb{R}^{n\times n} \to \mathbb{R}^{2n\times n}, \quad Du := \begin{pmatrix} Nu \\ uN^\top \end{pmatrix}$$

where

$$M = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

In order to apply the ADMM algorithm, we introduce a new auxiliary variable $v \in \mathbb{R}^{2n\times n}$ and give the following constraint convex optimization problem

$$\begin{aligned} \min_{u\in\mathbb{R}^{n\times n}, v\in\mathbb{R}^{2n\times n}} & \|v\|_1 + \frac{\alpha}{2}\|D^\gamma u - w\|_F^2 + \frac{\beta}{2}\|u - u_0\|_F^2, \\ \text{s.t. } & v = Du \end{aligned} \tag{14}$$

Then, let $p \in \mathbb{R}^{2n\times n}$ be the Lagrangian multiplier for the linear constraint $v = Du$, we give the corresponding augmented Lagrangian function,

$$\mathcal{L}(u, v, p) = \|v\|_1 + \frac{\alpha}{2}\|D^\gamma u - w\|_F^2 + \frac{\beta}{2}\|u - u_0\|_F^2 + \langle p, Du - v\rangle + \frac{\delta}{2}\|Du - v\|_F^2$$

where $\delta > 0$ is the penalty parameter. Therefore, under the framework of ADMM, the minimization is divided into two separated subproblems.

For solving the v-subproblem, we utilize the shrinkage formula [48]. The solution is computed iteratively by

$$\begin{aligned} v^{k+1} &= \arg\min_v \|v\|_1 + \frac{\delta}{2}\left\|Du^k - v + \frac{p^k}{\delta}\right\|_F^2 \\ &= shrink(Du^k + \frac{p^k}{\delta}, \frac{1}{\delta}) \end{aligned} \tag{15}$$

where $shrink(x, y) = \max(\|x\|_F - y, 0) \cdot \frac{x}{\|x\|_F}$, obeying the convention $0 \cdot \frac{0}{0} = 0$.

With respect to the u-subproblem, the minimization is written as

$$u^{k+1} = \arg\min_u \frac{\alpha}{2}\|D^\gamma u - w\|_F^2 + \frac{\beta}{2}\|u - u_0\|_F^2 + \langle p^k, Du - v^{k+1}\rangle + \frac{\delta}{2}\|Du - v^{k+1}\|_F^2.$$

By the first-order optimality condition [50], we obtain the simplified linear system as follows,

$$\left(\alpha(D^{\gamma})^{\top}D^{\gamma} + \delta D^{\top}D + \beta I\right)u = \alpha(D^{\gamma})^{\top}w + D^{\top}(\delta v^{k+1} - p^{k}) + \beta u_0.$$

By some simple operations, we obtain the following Sylvester-like equation

$$\left(\alpha M^2 + \delta N^{\top}N + \beta I\right)u + u(\alpha M^2 + \delta N^{\top}N) = \alpha(D^{\gamma})^{\top}w + D^{\top}(\delta v^{k+1} - p^{k}) + \beta u_0. \tag{16}$$

Then, we apply the conjugate gradient (CG) method [51] to solve the equation. In conclusion, the whole algorithm for fusing the noisy source images is given as follows.

Algorithm 1. ADMM algorithm for solving (13)

1: Initialize u^0, v^0 and p^0; set the parameters $\alpha, \beta, \gamma, \delta$.
2: For $k = 1, 2, \ldots, N$, compute iteratively u^{k+1}, v^{k+1}, and p^{k+1} by
1) Compute v^{k+1} by the shrinkage formula (15);
2) u^{k+1} is obtained by the CG method for solving (16);
3) Update the multipliers:

$$p^{k+1} = p^{k} + \tau\delta(Du^{k+1} - v^{k+1}).$$

Finally, motivated by the works in [42, 52], we give a following theorem to show the convergence of Algorithm 1.

Theorem 4.1. *For the fixed parameter $\delta > 0$ and $\tau \in (0, \frac{\sqrt{5}+1}{2})$, the ADMM algorithm for solving the FTVL2 model* (14) *is convergent.*

Proof. In order to prove the convergence of Algorithm 1, we first transform the proposed model into the general constrained convex problem. Specially, we set two functions as follows,

$$f(u) = \frac{\alpha}{2}\|D^{\gamma}u - w\|_F^2 + \frac{\beta}{2}\|u - u_0\|_F^2, \quad g(y) = \|v\|_1.$$

Then with the linear constrained condition $Du - v = 0$, we can deduce that for fixed $\delta > 0$ and $\tau \in (0, \frac{\sqrt{5}+1}{2})$, Algorithm 1 is convergent. □

Note that although the u-subproblem is approximatively solved by the conjugate gradient (CG) method, Algorithm 1 is empirically convergent.

In addition, for fair comparison with the FTVL1 model, we also apply the classic ADMM algorithm to solve (7). The corresponding discrete minimization problem is given as follows:

$$\min_{u\in\mathbb{R}^{n\times n}} \|Du\|_1 + \alpha\|D^{\gamma}u - w\|_1 + \frac{\beta}{2}\|u - u_0\|_F^2. \tag{17}$$

By introducing two variables v and t, we transforms into the following constrained minimization problem

$$\begin{aligned} &\min_{u,v,t} \|v\|_1 + \alpha\|t\|_1 + \frac{\beta}{2}\|u - u_0\|_F^2, \\ &\text{s.t. } v = Du, t = D^{\gamma}u - w. \end{aligned} \tag{18}$$

Therefore, the augmented Lagrangian function of (18) is formulated as

$$\mathcal{L}(u, v, t, p_1, p_2) = \|v\|_1 + \alpha\|t\|_1 + \frac{\beta}{2}\|u - u_0\|_F^2 + \langle p_1, Du - v\rangle + \frac{\delta_1}{2}\|Du - v\|_F^2 + \langle p_2, D^\gamma u - w - t\rangle + \frac{\delta_2}{2}\|D^\gamma u - w - t\|_F^2$$

where p_1 and p_2 denotes the multipliers, δ_1 and δ_2 denotes the positive penalty parameters. Similarly, we apply the following algorithm to solve the FTVL1 model.

Algorithm 2. ADMM algorithm for solving (17)

1: Initialize u^0, v^0, t^0, p_1^0 and p_2^0; set the parameters $\alpha, \beta, \gamma, \delta_1, \delta_2$.
2: For $k = 1, 2, \ldots, N$, compute iteratively u^{k+1}, v^{k+1}, t^{k+1}, p_1^{k+1} and p_2^{k+1} by
1) Compute v^{k+1} and t^{k+1} by the shrinkage formulas

$$\begin{aligned} v^{k+1} &= shrink(Du^k + \frac{p_1^k}{\delta_1}, \frac{1}{\delta_1}); \\ t^{k+1} &= shrink(D^\gamma u^k - w + \frac{p_2^k}{\delta_2}, \frac{\alpha}{\delta_2}); \end{aligned}$$

2) u^{k+1} is obtained by the CG method;
3) Update the multipliers:

$$\begin{aligned} p_1^{k+1} &= p_1^k + \tau\delta_1(Du^{k+1} - v^{k+1}), \\ p_2^{k+1} &= p_2^k + \tau\delta_2(D^\gamma u^{k+1} - w - t^{k+1}). \end{aligned}$$

5 Experiments

In this section, we present several numerical experiments to illustrate the superior performance of the proposed method for fusing the noisy source images. All numerical experiments are performed under Windows 8 and Matlab (R2015b) running on a desktop with 3.40GHz Intel Core i3-2130 CPU and 4G RAM memory. In the experiments, all the pixel values of the test source images belong to the interval $[0, 255]$. In order to show the efficiency of the proposed method, we compare with some existing popular algorithm including the WFTV model [20] and the FSGF model [23]. For the FSGF model, we first denoise the source images and then fuse the denoised images.

In real applications, it is not easy to quantitatively measure the quality of the fused results due to the absence of the true original image. According to [53], for evaluating the quality of the fused results, we just consider two of all the objective fusion metrics. The mutual information based fusion metric [54, 55] is defined by

$$Q_{MI} = MI(u^*, f_1) + MI(u^*, f_2),$$

where u^* is the fused image, f_1 and f_2 are the reference images. The mutual information operation $MI(x, y)$ can be expressed as

$$MI(x, y) = H(x) + H(y) - H(x, y)$$

with

$$H(x) = -\sum_{x} p(x)\log_2 p(x)$$
$$H(x,y) = -\sum_{x,y} p(x,y)\log_2 p(x,y),$$

where $p(x)$ is the marginal probability distribution function and $p(x,y)$ is the joint probability distribution function. The higher MI values implies the better fused results. Moreover, by combining with the principal (maximum and minimum) moments of the image phase congruency, a new fusion metric [56, 57] is defined as a product,

$$Q_P = (P_p)^{\alpha}(P_M)^{\beta}(P_m)^{\gamma},$$

where p is the phase congruency, M is the maximum moment, and m is the minimum moment, $P_p = \max\{C^p_{f_1u^*}, C^p_{f_2u^*}, C^p_{su^*}\}$, $P_M = \max\{C^M_{f_1u^*}, C^M_{f_2u^*}, C^M_{su^*}\}$, $P_m = \max\{C^m_{f_1u^*}, C^m_{f_2u^*}, C^m_{su^*}\}$. Note that s is the maximum-select map. The parameters α, β and γ can be tuned according to the importance of the three components. The correlation coefficient between two sets x and y is given by

$$C^k_{xy} = \frac{\sigma^k_{xy} + C}{\sigma^k_x \sigma^k_y + C},$$

where $k \in \{p, M, m\}$, σ_x and σ_y are their respective standard variances, σ_{xy} are the covariance of x and y, $C > 0$ is a constant.

For the fair compasions, we set the same maximum iteration number as 10 in the FSGF model, the FTVL1 model and the FTVL2 model. In Algorithm 1 and 2, since the parameters α and β balance the trade-off between the TV regularization and the data-fitting terms, we manually tune them for obtaining the good fused results. Then we set $1 \le \gamma \le 2$ in the fractional-order gradient operator and $0 \le \delta \le 10^{-5}$. Furthermore, since the parameter τ controls the convergence speed, we set $\tau = 1.618$ which the ADMM algorithm 1 converges faster than $\tau = 1$. In addition, the iterative number of the conjugate gradient method is equal to 5 for solving the u-subproblem.

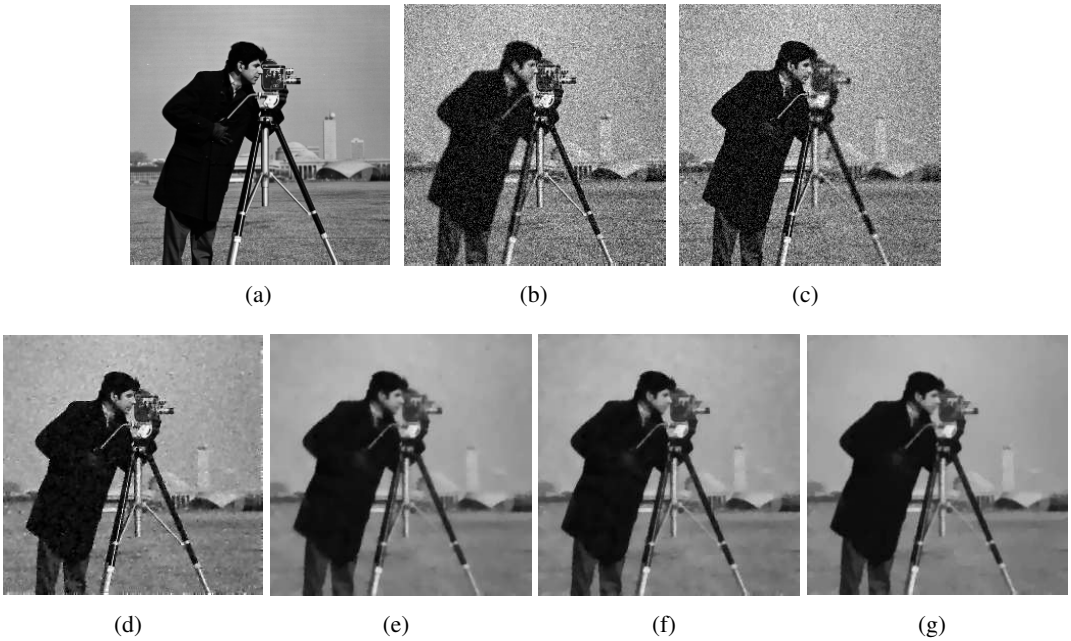
5.1 Experiment on the Image Cameraman

The test image "Cameraman" is blurred in the left part and right part respectively with a Gaussian kernel (mean zero, standard deviation $\sigma = 2$, kernel size 5×5). Then the Gaussian noise is added to the blurring images, where the noise level $\sigma \in \{15, 20, 25, 30\}$. We compare the fused results by using the FSGF, WFTV, FTVL1 and FTVL2 model. Table 1 lists the values of Q_{MI} and Q_P by using four different methods for the test image "Cameraman". By comparing the values of Q_{MI} and Q_P, the FTVL1 and FTVL2 model obtain the higher results than the FSGF and WFTV model. Moreover, the FTVL2 model obviously improves the values of Q_{MI}.

Figure 2 shows the fused images by using different methods for the test image "Cameraman", where the noise level $\sigma = 30$. We find that the noise is not thoroughly removed by the FSGF model, but the WFTV, FTVL1 and FTVL2 model are able to eliminate the white Gaussian noise in the source images (b) and (c). Unfortunately, the WFTV model suffers from some undesired false edges. The FTVL1 model products some artifacts in the background. The fused image by using the FTVL2 model is obviously better than the the WFTV and FTVL1 model. The FTVL2 model not only efficiently removes the white Gaussian noise, but also fuses highly the important features from two source images.

Table 1. The values of Q_{MI} and Q_P by using different methods for the test image "Cameraman" .

	Q_{MI}				Q_P			
Noise	FSGF	WFTV	FTVL1	FTVL2	FSGF	WFTV	FTVL1	FTVL2
15	5.4948	5.8548	5.9135	6.0294	0.5060	0.4955	0.5132	0.5195
20	5.1070	5.6269	5.6509	5.8047	0.4064	0.4353	0.4398	0.4490
25	4.7996	5.4410	5.4244	5.6272	0.3451	0.3935	0.3828	0.4007
30	4.5384	5.2760	5.2344	5.4845	0.2985	0.3583	0.3411	0.3672



(a) (b) (c)

(d) (e) (f) (g)

Figure 2. Comparison of different methods for fusing and denoising, the noise standard variance $\sigma = 30$. (a) The original image; (b) and (c) two source images by blurring the left part and right part; (d) FSGF [23]; (e) WFTV [20]; (f) FTVL1; (g) FTVL2.

For illustrating the superior performance of the proposed model, we give the corresponding edge maps of the fused images by the canny edge detector in Figure 3. By the edge maps comparison of the fused images, the FTVL2 model outperforms the three other methods. In the edge map of the FSGF model, there still exists much noise. However, although the region of the grass is smoothed by the WFTV, FTVL1 and FTVL2 model, there are no much artifacts in the region of the sky and the black coat for the FTVL2 model.

5.2 Experiment on the Real Image

In this section, we do several experiments on the real images for simultaneously denoising and fusing. Figure 4 shows the test real images, including three pairs of multi-focus images and five pairs of multi-modal images, which can be found in the website `http://www.quxiaobo.org/index.html`. The noisy images are obtained by adding the white Gaussian noise with the standard variance $\sigma \in \{15, 20, 25, 30\}$, respectively. First, we list Table 2 shown the values of Q_{MI} and Q_P by using four different model for the real multi-focus images. Our proposed FTVL2 model obtains the highest values of Q_{MI} by comparing with the FSGF, WFTV and FTVL1 model.

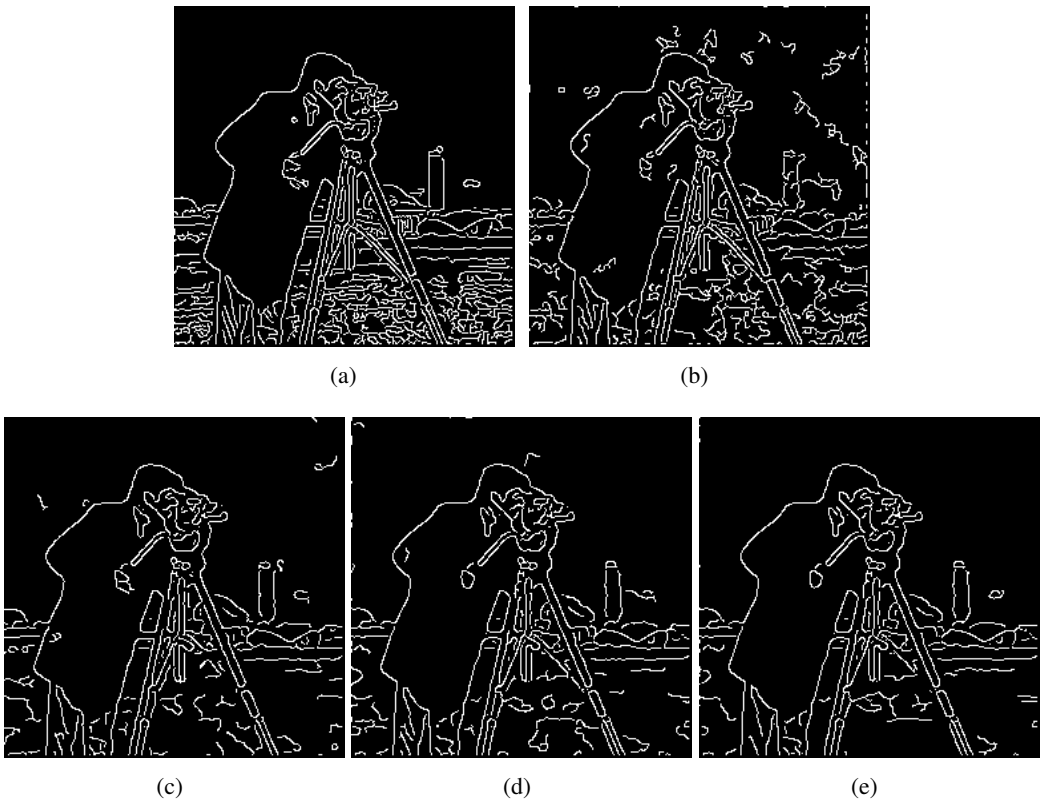


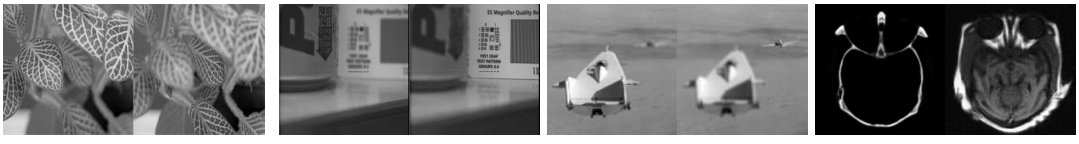
Figure 3. Edge maps of the fused images in Figure 2 by using the canny edge detector. (a) The original image; (b) FSGF [23]; (c) WFTV [20]; (d) FTVL1; (f) FTVL2.

But, by comparing with the values of Q_P, we nearly get the higher values than the other models. Especially, when the noise level is fairly large, the results of the proposed model shows the notable improvement.

Table 2. The values of Q_{MI} and Q_P by using different methods for the real multi-focus images.

		Q_{MI}				Q_P			
Image	Noise	FSGF	WFTV	FTVL1	FTVL2	FSGF	WFTV	FTVL1	FTVL2
Leaves	15	4.0347	3.9987	4.1202	4.1370	0.3192	0.3299	0.3637	0.3634
	20	3.8820	3.8466	3.9819	4.0141	0.2784	0.2762	0.2956	0.2965
	25	3.7433	3.7377	3.8746	3.9069	0.2382	0.2370	0.2430	0.2476
	30	3.6106	3.6350	3.7671	3.8061	0.2075	0.2086	0.2032	0.2129
Book	15	5.3383	6.0771	6.0383	6.0930	0.2698	0.3343	0.3228	0.3401
	20	5.0130	5.8697	5.8207	5.9102	0.2072	0.2952	0.2868	0.2954
	25	4.7648	5.6822	5.6342	5.7364	0.1678	0.2649	0.2516	0.2659
	30	4.5413	5.5135	5.4604	5.5848	0.1401	0.2415	0.2265	0.2514
Plane	15	4.8575	5.6256	5.6138	5.6456	0.2875	0.3722	0.3503	0.3499
	20	4.4670	5.3921	5.3303	5.3998	0.2372	0.3335	0.3007	0.3107
	25	4.1008	5.1903	5.0720	5.2062	0.1939	0.3022	0.2641	0.2864
	30	3.8178	5.0092	4.8324	5.0186	0.1573	0.2754	0.2350	0.2519

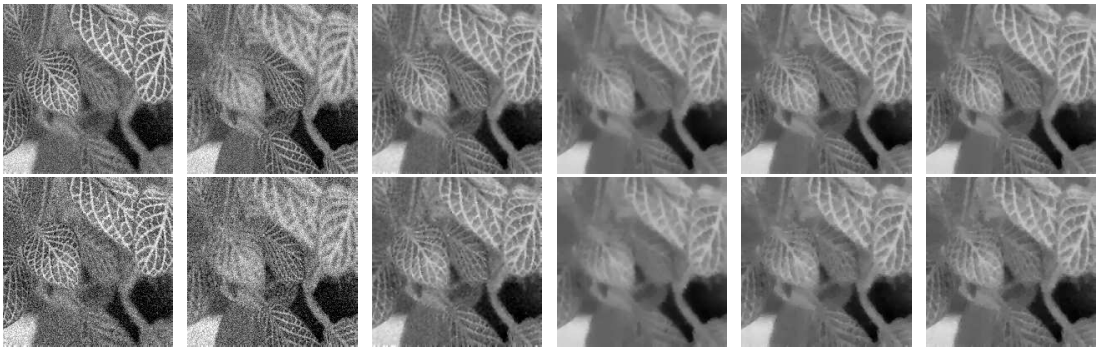
Figure 5 and 7 show the fused images by applying four different models, where the white Gaussian noise $\sigma = 20, 30$. Since the FSGF model is just used to fuse the noiseless source images,



(a) Leaves (204×204) (b) Book (256×256) (c) Plane (160×160) (d) Brain (256×256)

(e) Crossroad (256×256) (f) Hill (240×240) (g) Grass (256×256) (h) Plant (256×256)

Figure 4. Source images.



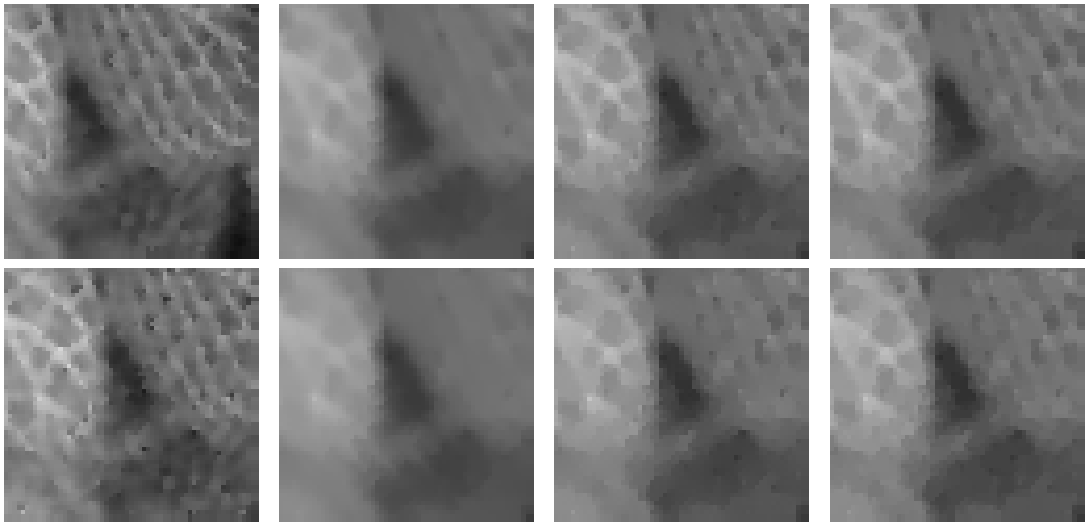
(a) (b) (c) (d) (e) (f)

Figure 5. Comparison of different models for the images "Leaves", where the white Gaussian noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) and (b): two input source images; (c) FSGF [23]; (d) WFTV [20]; (e) FTVL1; (f) FTVL2.

we find that the fused result (c) still retains some noise. The WFTV model is able to remove the noise, but oversmoothes the details of images. Compared with the FSGF and WFTV model, the fused images of the FTVL1 and FTVL2 model are obviously improved. Furthermore, the fused images of the FTVL2 model preserve a little clearer texture than the FTVL1 model. In conclusion, the proposed FTVL2 model can effectively remove the white Gaussian noise as well as fuse the salient feature from the source images. Especially, we give the zoomed parts of the fused images by using different models, shown in Figure 6. The texture of the leaves by the FTVL2 model is clearer than the FTVL1 model.

Secondly, we do the experiments for the test multi-modal images by using four different models. Table 3 shows the values of Q_{MI} and Q_P by applying four different models, where the noise level $\sigma = 15, 20, 25, 30$. The FTVL1 and FTVL2 model almost obtain the higher values with respect to Q_{MI} and Q_P than the FSGF and WFTV model. However, the fused results of the proposed FTVL2 model is better than the FTVL1 model. Due to this reason, we observe that the FTVL2 model is more suitable for fusing the noisy source images.

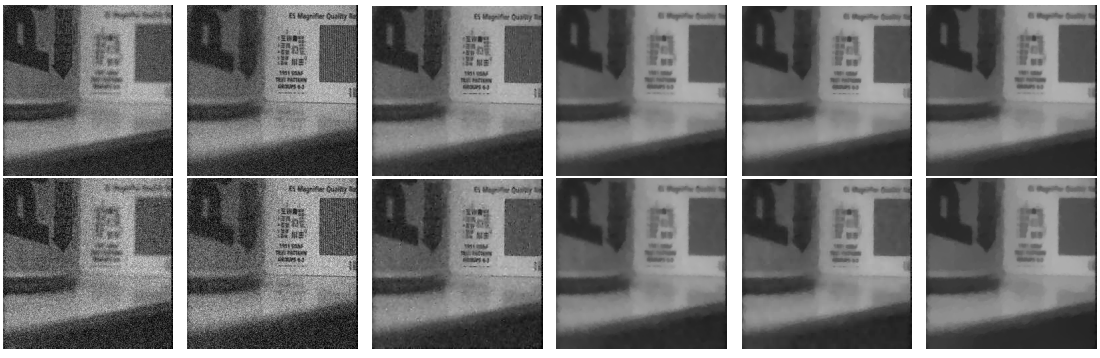
Figure 8 shows the Comparison of different models for fusing the noisy computed tomography (CT) image and the noisy magnetic resonance imaging (MRI) image, where the white Gaussian noise $\sigma = 20$ and $\sigma = 30$, respectively. Figure 8(a) shows the CT images which well provides the detail structure of the body's bone, Figure 8(b) shows the MRI images which mainly illuminates the detail structure of the body's soft tissue. Figure 8(c-f) show the fused results by applying the



(a) (b) (c) (d)

Figure 6. Comparison of zoomed fused images in Figure 5, where the white Gaussian noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) FSGF [23]; (b) WFTV [20]; (c) FTVL1; (d) FTVL2.

FSGF model [23],the WFTV model [20], the FTVL1 model and the FTVL2 model. Figure 8(c) is the fused result of the FSGF model, there still retains some noise. But from Figure 8(d-f), the noise is effectively removed by the WFTV model [20], the FTVL1 model and the FTVL2 model. Unfortunately, the WFTV model darken the fused image and reduce the image contrast. The FTVL1 model and the FTVL2 model gets the better fused images. But the FTVL1 model products the undesired artifacts in the fused images. The FTVL2 model fuses the abundant information and salient feature of the noisy source images, especially by comparing with the zoomed parts of the fused image in Figure 9. By compared with the FTVL1 model, the fused images of the FTVL2 model better fit with the quality perception of the human visual system (HVS). Therefore, the proposed model effectively removes the noise as well as fuses the salient detail of the CT images and MRI images.



(a) (b) (c) (d) (e) (f)

Figure 7. Comparison of different models for the images "Book", where the white Gaussian noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) and (b): two input source images; (c) FSGF [23]; (d) WFTV [20]; (e) FTVL1; (f) FTVL2.

Table 3. The values of Q_{MI} and Q_P by using different models for the real multi-modal images.

		Q_{MI}				Q_P			
Image	Noise	FSGF	WFTV	FTVL1	FTVL2	FSGF	WFTV	FTVL1	FTVL2
Brain	15	2.6017	2.5786	2.7504	2.7885	0.2057	0.2014	0.2192	0.2133
	20	2.4345	2.4550	2.6302	2.6861	0.1553	0.1698	0.1788	0.1735
	25	2.3198	2.3498	2.5310	2.6105	0.1268	0.1481	0.1496	0.1529
	30	2.2203	2.2695	2.4614	2.5499	0.1035	0.1314	0.1285	0.1385
Grass	15	3.1104	2.6442	3.1401	3.1529	0.4861	0.3601	0.4891	0.4943
	20	2.9784	2.5556	2.9814	2.9981	0.4227	0.2943	0.4123	0.4208
	25	2.8489	2.4823	2.8461	2.8892	0.3602	0.2449	0.3448	0.3586
	30	2.7402	2.4173	2.7420	2.7936	0.3079	0.2094	0.2894	0.3093
Crossroad	15	2.7077	2.1029	2.8007	2.8271	0.2774	0.2531	0.2958	0.3012
	20	2.4283	1.9897	2.5829	2.6417	0.2132	0.2185	0.2481	0.2492
	25	2.2137	1.9043	2.4015	2.4753	0.1669	0.1916	0.2124	0.2147
	30	2.0501	1.8284	2.2745	2.3578	0.1388	0.1709	0.1813	0.1901
Hill	15	2.5245	1.7110	2.5897	2.9196	0.1073	0.1091	0.1105	0.01147
	20	2.3027	1.6892	2.4380	2.4840	0.0797	0.0925	0.0901	0.0926
	25	2.1320	1.6677	2.3204	2.3704	0.0611	0.0806	0.0748	0.0806
	30	2.0114	1.6477	2.2236	2.2957	0.0516	0.0724	0.0656	0.0737
Plant	15	3.1427	3.1157	3.3199	3.3327	0.3800	0.3734	0.3935	0.4000
	20	3.1056	2.9502	3.1316	3.1385	0.3175	0.3122	0.3117	0.3194
	25	2.9366	2.8258	2.9659	2.9919	0.2527	0.2553	0.2531	0.2639
	30	2.8034	2.7230	2.8278	2.8532	0.2137	0.2197	0.2083	0.2208

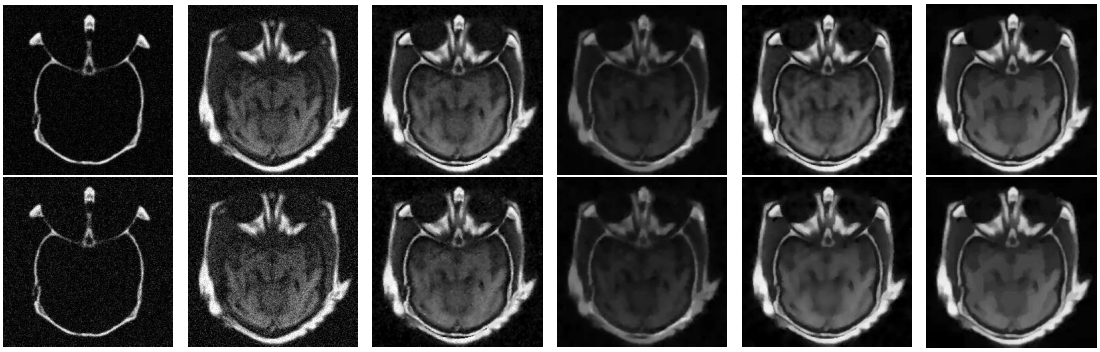
Figure 10 and 11 are the Comparison of different models for fusing the noisy visible image and the noisy infrared image. The visible image captures many spatial details and background information, while the infrared image easily determines the position of thermal objects. The visible image is different from the infrared image. And the information of the visible image and the infrared image is complementary. Obviously, for the FSGF model, the noise is not thoroughly removed. The WFTV model reduce the contrast of the fused image. But, for the proposed FTVL1 and FTVL2 models, all the salient features of are fairly clear. Furthermore, the proposed FTVL2 model obtains the better fused image without the undesired artifacts.

6 Conclusion

In this paper, according to the underlying properties of the fractional-order derivative, we assume that the synthetical fractional-order gradient information from the source images fits with the fractional-order gradient information of the fused image in the sense of L^2 norm. By combining the total variation (TV) regularization for removing the noise with the fidelity term for making the fused image close a predefined image, we present a novel convex variational model for fusing the noisy source images. Then, an alternating direction method of multiplier (ADMM) is developed for solving the proposed FTVL2 model. Numerical experiments show that our method outperform the state-of-the-art variational methods in terms of visual and quantitative measures.

Acknowlegenments

We would like to thank Fang Li and Weiwei Wang for providing the software codes [23] and [20].



(a) (b) (c) (d) (e) (f)

Figure 8. Comparison of different models for fusing and denoising in the images "Brain", where white Gaussian noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) and (b): two source images; (c) FSGF [23]; (d) WFTV [20]; (e) FTVL1; (f) FTVL2.

References

[1] R. S. Blum, *Multi-Sensor Image Fusion and its Applications, (special series on Signal Processing and Communications)*. 2005.

[2] N. Burgos, M. J. Cardoso, M. Modat, S. Pedemonte, J. Dickson, A. Barnes, J. S. Duncan, D. Atkinson, S. R. Arridge, B. F. Hutton, and S. Ourselin, *Attenuation Correction Synthesis for Hybrid PET-MR Scanners*, pp. 147–154. Springer Berlin Heidelberg, 2013.

[3] A. Ardeshir Goshtasby and S. Nikolov, "Guest editorial: Image fusion: Advances in the state of the art," *Information Fusion*, vol. 8, no. 2, pp. 114–118, 2007.

[4] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information Fusion*, vol. 13, no. 1, pp. 10–19, 2012.

[5] V. Tsagaris and V. Anastassopoulos, "Information measure for assessing pixel-level fusion methods," in *Proceedings of SPIE*, vol. 5573, pp. 64–71, 2004.

[6] N. Mitianoudis and T. Stathaki, "Pixel-based and region-based image fusion schemes using ICA bases," *Information Fusion*, vol. 8, no. 2, pp. 131–142, 2007.

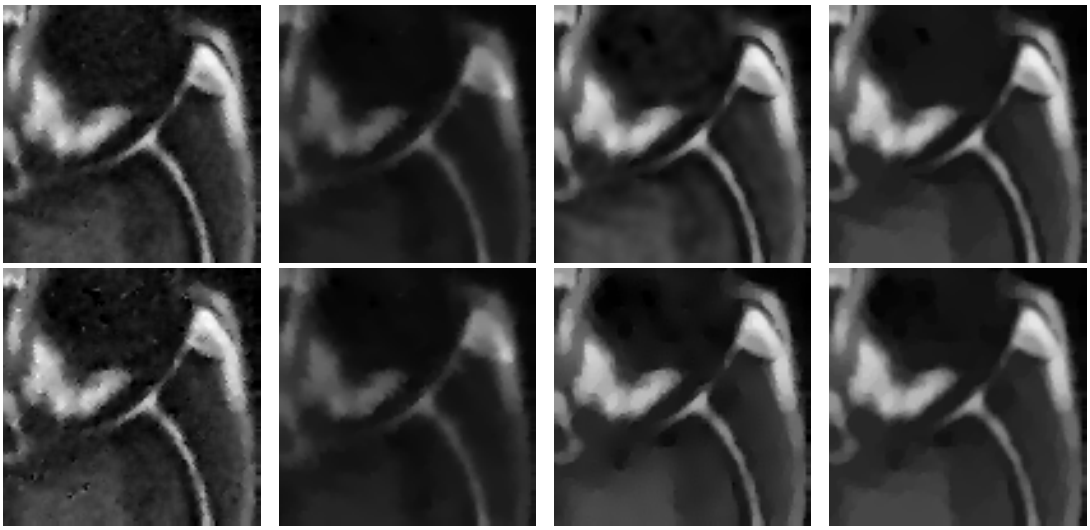
[7] M. Kumar and S. Dass, "A total variation-based algorithm for pixel-level image fusion," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2137–2143, 2009.

[8] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image and Vision Computing*, vol. 26, no. 7, pp. 971–979, 2008.

[9] P. Shah, S. N. Merchant, and U. B. Desai, "An efficient spatial domain fusion scheme for multifocus images using statistical properties of neighborhood," in *2011 IEEE International Conference on Multimedia and Expo*, pp. 1–6, 2011.

[10] S. Li, J. T. Kwok, and Y. Wang, "Combination of images with diverse focuses using the spatial frequency," *Information Fusion*, vol. 2, no. 3, pp. 169–176, 2001.

[11] G. Piella, "Image fusion for enhanced visualization: A variational approach," *International Journal of Computer Vision*, vol. 83, no. 1, pp. 1–11, 2009.



(a) (b) (c) (d)

Figure 9. Comparison of zoomed fused images in Figure 8, where the white Gaussian noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) FSGF [23]; (b) WFTV [20]; (c) FTVL1; (d) FTVL2.

[12] D. A. Socolinsky and L. B. Wolff, "Multispectral image visualization through first-order fusion," *IEEE Transactions on Image Processing*, vol. 11, no. 8, pp. 923–931, 2002.

[13] A. Toet, "A morphological pyramidal image decomposition," *Pattern Recognition Letters*, vol. 9, no. 4, pp. 255–261, 1989.

[14] V. S. Petrovic and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228–237, 2004.

[15] H. Li, B. S. Manjunath, and S. K. Mitra, "Multi-sensor image fusion using the wavelet transform," in *Proceedings of 1st International Conference on Image Processing*, vol. 1, pp. 51–55, 1994.

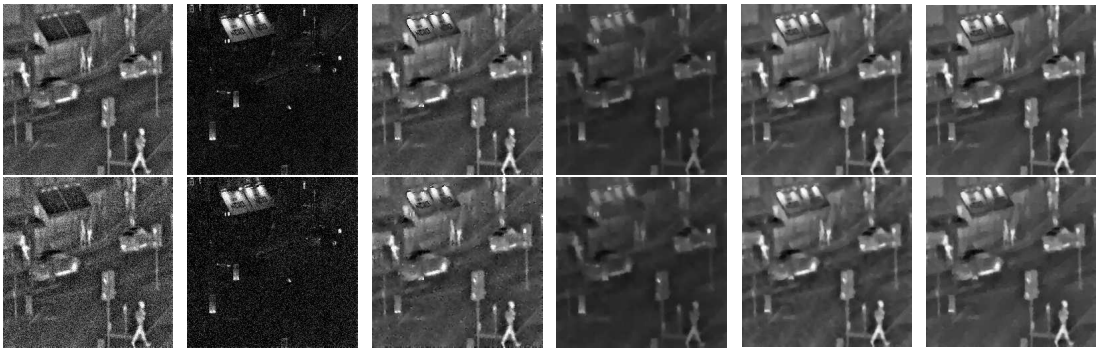
[16] V. Naidu and J. Raol, "Pixel-level image fusion using wavelets and principal component analysis," *Defence Science Journal*, vol. 58, no. 3, p. 338, 2008.

[17] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, "Remote sensing image fusion using the curvelet transform," *Information Fusion*, vol. 8, no. 2, pp. 143–156, 2007.

[18] Y. Chai, H. Li, and X. Zhang, "Multifocus image fusion based on features contrast of multiscale products in nonsubsampled contourlet transform domain," *Optik-International Journal for Light and Electron Optics*, vol. 123, no. 7, pp. 569–581, 2012.

[19] G. Guorong, X. Luping, and F. Dongzhu, "Multi-focus image fusion based on nonsubsampled shearlet transform," *IET Image Processing*, vol. 7, pp. 633–639, 2013.

[20] W. W. Wang, P. L. Shui, and X. C. Feng, "Variational models for fusion and denoising of multifocus images," *IEEE Signal Processing Letters*, vol. 15, pp. 65–68, 2008.



(a) (b) (c) (d) (e) (f)

Figure 10. Comparison of different method for fusing and denoising in the image "Crossroad" with Gaussian white noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) and (b) two source images; (c)-(f) fused images by applying different methods: (c) FSGF [23]; (d) WFTV [20]; (e) FTVL1; (f) FTVL2.

(a) (b) (c) (d) (e) (f)

Figure 11. Comparison of different method for fusing and denoising in the image "Hill" with Gaussian white noise $\sigma = 20$ (in the 1st row) and $\sigma = 30$ (in the 2nd row), respectively. (a) and (b) two source images; (c)-(f) fused images by applying different methods: (c) FSGF [23]; (d) WFTV [20]; (e) FTVL1; (f) FTVL2.

[21] B. Lu, C. Miao, and H. Wang, "Pixel level image fusion based on linear structure tensor," in *2010 IEEE Youth Conference on Information, Computing and Telecommunications*, pp. 303–306, 2010.

[22] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Information Fusion*, vol. 20, pp. 60 – 72, 2014.

[23] F. Li and T. Zeng, "Variational image fusion with first and second-order gradient information," *Journal of Computational Mathematics*, vol. 34, no. 2, pp. 200–222, 2016.

[24] Y. Xu, T. Z. Huang, J. Liu, and X. G. Lv, "Split Bregman iteration algorithm for image deblurring using fourth-order total bounded variation regularization model," *Journal of Applied Mathematics*, vol. 2013, no. 3, pp. 417–433, 2013.

[25] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[26] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Modeling & Simulation*, vol. 8, no. 2, pp. 337–369, 2010.

[27] L. Yu and W. Zengfu, "Simultaneous image fusion and denoising with adaptive sparse representation," *IET Image Processing*, vol. 9, pp. 347–357(10), 2015.

[28] J. Zhang and K. Chen, "A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2487–2518, 2015.

[29] I. Podlubny, *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, vol. 198. Academic press, 1998.

[30] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*. Wiley-Interscience, 1 ed., 1993.

[31] K. Oldham and J. Spanier, *The fractional calculus theory and applications of differentiation and integration to arbitrary order*, vol. 111. Elsevier, 1974.

[32] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.

[33] B. He and X. Yuan, "On the O(1/n) convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal on Numerical Analysis*, vol. 50, no. 2, pp. 700–709, 2012.

[34] C. Chen, M. K. Ng, and X.-L. Zhao, "Alternating direction method of multipliers for nonlinear image restoration problems," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 33–43, 2015.

[35] Z. Ren, C. He, and Q. Zhang, "Fractional order total variation regularization for image super-resolution," *Signal Processing*, vol. 93, no. 9, pp. 2408 – 2421, 2013.

[36] J. Zhang, Z. Wei, and L. Xiao, "A fast adaptive reweighted residual-feedback iterative algorithm for fractional-order total variation regularized multiplicative noise removal of partly-textured images," *Signal Processing*, vol. 98, pp. 381 – 395, 2014.

[37] Y. Zhang, W. Zhang, Y. Lei, and J. Zhou, "Few-view image reconstruction with fractional-order total variation," *J. Opt. Soc. Am. A*, vol. 31, no. 5, pp. 981–995, 2014.

[38] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations, Volume 204 (North-Holland Mathematics Studies)*. New York, NY, USA: Elsevier Science Inc., 2006.

[39] I. Podlubny, A. Chechkin, T. Skovranek, Y. Chen, and B. M. V. Jara, "Matrix approach to discrete fractional calculus ii: Partial fractional differential equations," *Journal of Computational Physics*, vol. 228, no. 8, pp. 3137 – 3153, 2009.

[40] H. Wang and N. Du, "Fast solution methods for space-fractional diffusion equations," *Journal of Computational and Applied Mathematics*, vol. 255, pp. 376 – 383, 2014.

[41] J.-J. Mei, Y. Dong, T.-Z. Huang, and W. Yin, "Cauchy noise removal by nonconvex admm with convergence guarantees," *Journal of Scientific Computing (online)*, May 2017.

[42] B. He and H. Yang, "Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities," *Operations Research Letters*, vol. 23, no. 35, pp. 151 – 161, 1998.

[43] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[44] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognition*, vol. 37, no. 9, pp. 1855 – 1872, 2004.

[45] Q. Zhang and B. long Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Processing*, vol. 89, no. 7, pp. 1334 – 1346, 2009.

[46] J.-J. Mei and T.-Z. Huang, "Primal-dual splitting method for high-order model with application to image restoration," *Applied Mathematical Modelling*, vol. 40, no. 3, pp. 2322 – 2332, 2016.

[47] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems withapplications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[48] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[49] T.-H. Ma, T.-Z. Huang, X.-L. Zhao, and Y. Lou, "Image deblurring with an inaccurate blur kernel using a group-based low-rank image prior," *Information Sciences*, vol. 408, pp. 213 – 233, 2017.

[50] M. K. Ng, F. Wang, and X. Yuan, "Inexact alternating direction methods for image recovery," *SIAM Journal on Scientific Computing*, vol. 33, no. 4, pp. 1643–1668, 2011.

[51] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, vol. 49. Journal of Research of the National Bureau of Standards, 1952.

[52] Glowinski and Roland, *Numerical Methods for Nonlinear Variational Problems*. Springer-Verlag, 1984.

[53] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 94–109, Jan 2012.

[54] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.

[55] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on'information measure for performance of image fusion'," *Electronics letters*, vol. 44, no. 18, pp. 1066–1067, 2008.

[56] J. Zhao, R. Laganiere, and Z. Liu, "Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement," *International Journal of Innovative Computing, Information and Control*, vol. 3, no. 6, pp. 1433–1447, 2007.

[57] Z. Liu, D. S. Forsyth, and R. Laganiere, "A feature-based metric for the quantitative evaluation of pixel-level image fusion," *Computer Vision and Image Understanding*, vol. 109, no. 1, pp. 56 – 68, 2008.