Journal of Computational Mathematics
Vol.34, No.2, 2016, 200–222.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1512-m2014-0008

VARIATIONAL IMAGE FUSION WITH FIRST AND SECOND-ORDER GRADIENT INFORMATION*

Fang Li
Department of Mathematics, East China Normal University, and Shanghai Key Laboratory of Pure Mathematics and Mathematical Practice, Shanghai, China
Email: fli@math.ecnu.edu.cn
Tieyong Zeng
Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
Email: zeng@hkbu.edu.hk

Abstract

Image fusion is important in computer vision where the main goal is to integrate several sources images of the same scene into a more informative image. In this paper, we propose a variational image fusion method based on the first and second-order gradient information. Firstly, we select the target first-order and second-order gradient information from the source images by a new and simple salience criterion. Then we build our model by requiring that the first-order and second-order gradient information of the fused image match with the target gradient information, and meanwhile the fused image is close to the source images. Theoretically, we can prove that our variational model has a unique minimizer. In the numerical implementation, we take use of the split Bregman method to get an efficient algorithm. Moreover, four-direction difference scheme is proposed to discrete gradient operator, which can dramatically enhance the fusion quality. A number of experiments and comparisons with some popular existing methods demonstrate that the proposed model is promising in various image fusion applications.

Mathematics subject classification: 65N06, 65B99.
Key words: Image fusion, Feature selection, Bounded variation, Second bounded variation, Split Bregman

1. Introduction

Image fusion has become an active issue in image processing and computer vision owing to the availability of multisensor data in many fields. The main goal of image fusion is to integrate multiple source images of the same scene into a single highly informative image which is more suitable for human or computer vision. Despite the dissimilarity of the source images, they are highly correlated with each other and complementary in nature [17]. Hence, by fusing these images into a single one, remarkable improvement can be expected. In recent years, image fusion has attracted a large amount of attention in a wide variety of applications such as concealed weapon detection, remote sensing, medical diagnosis, defect inspection, and military surveillance [4, 8]. For example, due to limited depth-of-focus of optical lenses in CCD devices, it is extremely hard to get an ideal image such that all the involved objects are "in focus". However, in each image, only the object in focus is clear. Image fusion process is thus required

* Received July 21, 2014 / Revised version received August 18, 2015 / Accepted December 29, 2015 / Published online March 6, 2016 /

to provide an all "in focus" ideal image [34]. In medical imaging, computer tomography (CT) is usually good for imaging bone structures, while magnetic resonance imaging (MRI) is more suitable for soft tissues. Hence, by fusion, it is possible to obtain a single image that describes bone structure as well as soft tissues, which evidently is important for medical diagnosis [36]. In remote sensing imagery, some sensors are good at capturing high resolution spatial information without spectral information, while some other sensors are powerful at recording spectral information but with low spatial information. Fusing these different types of data thus could provide images with both spectral information and high resolution spatial information [2, 32].

During the past two decades, many image fusion methods have been developed. Based on the levels where information is integrated, image fusion methods can be roughly classified into three levels: pixel-level, feature-level and decision-level [1]. In pixel-level fusion, the pixel values of the fused image are derived from the pixel values of all the source images by some principles [38]. The main advantage of pixel-level fusion is that the original measured quantities are directly processed and the algorithms are easy to implement [35]. We focus on pixel-level fusion in this paper. Moreover, generally speaking, the pixel-level fusion methods can be categorized into spatial domain fusion methods [23,25,26,45] and transform domain fusion methods [9,29,30,34]. Let us first introduce some notations before reviewing the fusion methods.

Assume that $f_i : \Omega \to \mathbb{R}, i = 1, \cdots, m$ are the source images to be fused, where $\Omega \subset \mathbb{R}^2$ denotes the image domain which typically is a bounded rectangle. For each pixel $x \in \Omega$, the value $f_i(x)$ represents the gray level at x. Furthermore, let us suppose that $u : \Omega \to \mathbb{R}$ is the required fused image. Assuming that $\mathcal{S}(\cdot)$ represents certain feature selection rule, the space domain fusion can be formulated as [44, chapter 12]:

$$u(x) = \mathcal{S}\left(f_1(x), \cdots, f_m(x)\right).$$

Simple examples of spatial domain fusion include average and weighted average of the source images [26, 45]. On the other side, transform domain fusion method enables the use of a rather general framework where the salient image features are more clearly depicted than in the spatial domain. Let $\mathcal{T}$ denote a transform operator and $\mathcal{S}(\cdot)$ again the feature selection rule. The transform domain fusion method can be outlined as:

$$u(x) = \mathcal{T}^{-1}\Big\{\mathcal{S}\left\{\mathcal{T}\left(f_1(x), \cdots, f_m(x)\right)\right\}\Big\}.$$

The transform domain based method has been popular ever since the introduction of pyramid transform in mid-80's [9]. For instance, we have the Laplacian pyramid [9] and Gradient pyramid [10]. As the development of wavelet methods, wavelet multiscale decomposition is later applied to replace pyramid decomposition in image fusion with the similar idea. Evidently, the key step in transform domain fusion is the selection rule of transform coefficient. Many strategies have been developed in literatures [10, 29, 34, 49].

Recently, some variational fusion methods [36, 41, 43] have been introduced based on structure tensor and first-order gradient information. Structure tensor is widely used to enhance coherence in image restoration problems [15, 27, 47]. Since the first-order method is closely related to the proposed method in this paper, we recall the basic idea of first-order fusion approach below (for details, see [36]).

Let us first introduce the structure tensor which is widely used to extract local image feature from the source images. Giving source images $f_i : \Omega \to \mathbb{R}, i = 1, \cdots, m$, the structure tensor is defined as

$$G(x) := \sum_{i=1}^{m} \nabla f_i(x) \nabla f_i^T(x), \quad \forall x \in \Omega,$$

where $\nabla f_i := (\frac{\partial f_i}{\partial x}, \frac{\partial f_i}{\partial y})$. Note that for each x , G is a symmetric matrix which can be decomposed as

$$G = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T,$$

where λ_1, λ_2 denote the largest and smallest eigenvalues of G, and $\mathbf{e}_1, \mathbf{e}_2$ denote their corresponding unit eigenvectors. Geometrically, the first eigenvector $\mathbf{e}_1$ indicates the orientation maximizing the gray-value fluctuations, while $\mathbf{e}_2$ is perpendicular to $\mathbf{e}_1$. The eigenvalues λ_1, λ_2 convey shape information. In practice, one can often use a smoothed version of f_i in the construction of structure tensor to remove noise and fine structures. Remark that a weighted version of structure tensor is used in [36] and structure tensor in multiple resolution is used in [43] for image fusion. Basically, in the first-order fusion method, the target vector field $V : \Omega \to \mathbb{R}^2$ of the fused image u is constructed as:

$$V := \sqrt{\lambda_1} \mathbf{e}_1,$$

where the direction of $\mathbf{e}_1$ is chosen to satisfy $\langle \mathbf{e}_1, \sum_{i=1}^m \nabla f_i \rangle \geq 0$. Then the fused image u is obtained by solving the following problem:

$$\min_u \int_\Omega |\nabla u - V|^2 dx. \tag{1.1}$$

In [36], an improved version of first-order method (1.1) is proposed by adding some other variational terms from [3] to enhance the local contrast. Generally gradient descent method is used in the numerical implementation. Model (1.1) is effective in image fusion, however, there is still room for improvement.

In this paper, we consider a somewhat different approach as a complementary. Indeed, firstly, we construct other target information from the source images which can preserve the desired prominent information. Since higher order methods have been widely used and proven to be very effective in image society [5,7,13,22,31], we intend to use both the first-order and the second-order gradient information as the target. Meanwhile, a new salience criterion is built to choose the target information from the given source images. Secondly, we propose a new model based on (1.1). In the proposed model, we require that the corresponding gradient information of fused image match with the target first-order and second-order gradient information in the sense of L^1 norm, meanwhile the fused image is close to the source images in the sense of L^2 norm. Theoretically, the existence of a unique minimizer for the proposed model is proved. Numerically, a fast algorithm is proposed using split Bregman technique and an effective four-direction difference scheme is proposed to discrete the gradient operator.

The paper is organized as follows. In Section 2, we present the salience criterion for feature selection, and then build the novel variational model with a theoretical analysis of the existence of the minimizer. Then in Section 3 we give the fast algorithm based on the split Bregman technique. The numerical experiments are reported in Section 4. Finally, we conclude the paper in Section 5.

2. The Proposed Model

2.1. Feature selection

Assume that the source images are $f_i, i = 1, \cdots, m$, which are almost noise-free and aligned. If the images are noisy or not aligned, we need do denoising or registration processing first. Since image denoising and registration are widely studied topics in literature and many effective methods are proposed, we will not go to the details. Now suppose that we have the first-order and second-order gradient fields of the source images which are represented by ∇f_i and $\nabla^2 f_i, i = 1, \cdots, m$ respectively.

Our first task is to construct the target first-order and second-order gradient fields based on the source images. Our basic idea is to choose the most salient feature in each component of the gradient information. Inspired by the feature selection strategies in literatures [10, 29, 34, 49], here we propose a very simple and effective one which is then integrated into our coming variational model for image fusion. We remark that other feature selection rule can also be used in this step.

Assume that we have two features $M_i : \Omega \to \mathbb{R}, i = 1, 2$. For $x, y \in \Omega$, let us define a mean kernel function

$$K(x, y) = \begin{cases} 1/|\omega|, & \text{if } y \in \omega_x, \\ 0, & \text{otherwise}, \end{cases}$$

where ω_x denotes a bounded neighborhood of the pixel x with area $|\omega|$. Using the average value of M_i^2 as a salience measure, a binary mask can be obtained by choosing the salient feature as follows,

$$B(x) = \begin{cases} 1, & \text{if } K * M_1^2 > K * M_2^2, \\ 0, & \text{otherwise}. \end{cases}$$

After that, the fusion binary mask is given by smoothing and thresholding

$$\widetilde{B}(x) = \begin{cases} 1, & \text{if } K * B > 0.5, \\ 0, & \text{otherwise}. \end{cases}$$

This step aims to make the selection more consistent in neighborhoods since isolated points or tiny structures will be removed. Finally, the selected feature is set to be

$$M = \widetilde{B} M_1 + (1 - \widetilde{B}) M_2.$$

The above selection scheme is designed for two features. In the case of more than two features, we use recursive method and at one time only two features are combined. Using this scheme, we can select first-order feature from each component of $\nabla f_1, \cdots, \nabla f_m$ and denote the result as $\mathbf{v}$. Similarly, we select second-order feature from $\nabla^2 f_1, \cdots, \nabla^2 f_m$ and denote the result as $\mathbf{w}$. By this procedure we get the target first-order feature $\mathbf{v}$ and the target second-order feature $\mathbf{w}$.

2.2. The proposed model

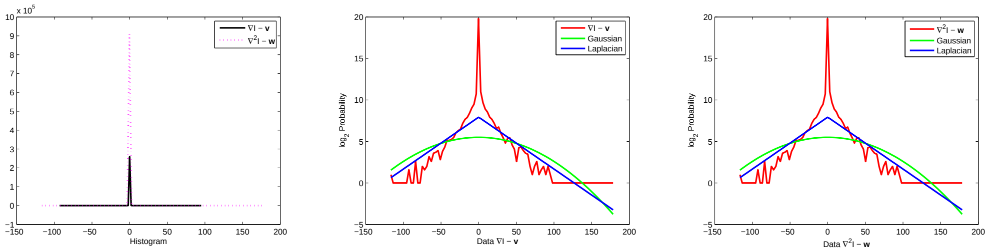
We are aim to search for the fused image which combines all the important salient features in the source images. So the fused image u is expected to have first-order gradient information matching with the target $\mathbf{v}$, and second-order gradient information matching with $\mathbf{w}$. Similar as functional (1.1), we can immediately write down the cost function:

$$\int_\Omega |\nabla u - \mathbf{v}|^2 + \alpha |\nabla^2 u - \mathbf{w}|^2 dx, \tag{2.1}$$

where $|\mathbf{s}|(x) := \sqrt{\sum \mathbf{s}_i(x)^2}$ denotes the Euclidean distance and α is a positive balancing parameter. In contrast, we prefer to use L^1 norm to measure the distance here:

$$\int_\Omega |\nabla u - \mathbf{v}| + \alpha |\nabla^2 u - \mathbf{w}| dx. \tag{2.2}$$

The underline reason to choose not L^2 but L^1 is as follows. Firstly, since images usually have sparse gradient information, that is, $\nabla u, \nabla f_i, \nabla^2 u, \nabla^2 f_i$ are all sparse. Then we can also assume that $\mathbf{v}$ and $\mathbf{w}$ are sparse by the selection rule. Hence $\nabla u - \mathbf{v}$ and $\nabla^2 u - \mathbf{w}$ are also sparse. As is well known, L^1 norm ensures more sparsity than L^2 norm. Secondly, in statistics, L^2 norm comes from Gaussian distribution, while L^1 norm comes from Laplacian distribution by taking $-\log$ operation. To show the difference, we take the cameraman test images shown in the first row of Fig. 4.1 as an example (see Section 4). The desired fusion result is u equals the ground truth image I in Fig. 4.1(a). $\mathbf{v}$ and $\mathbf{w}$ are the selected first order and second order features from the source images Fig. 4.1(b)-(c) by method described in Section 2.1. We show the distribution of data $\nabla I - \mathbf{v}$ and $\nabla^2 I - \mathbf{w}$ in Fig. 2.1 together with the Gaussian and Laplacian fitting results. We find that the distributions of the first and second order gradient errors $\nabla I - \mathbf{v}$ and $\nabla^2 I - \mathbf{w}$ match Laplacian distribution better than Gaussian distribution. Indeed, Laplacian fitting has a mean square error (MSE) much smaller than Gaussian fitting. Hence, it is more reasonable to use L^1 norm to measure the distance as (2.2) than use L^2 norm as in (2.1).



(a) (b) (c)

Fig. 2.1. A Laplacian distribution is a better model of $\nabla I - \mathbf{v}$ and $\nabla^2 I - \mathbf{w}$ than a Gaussian distribution. (a) The histogram of $\nabla I - \mathbf{v}$ and $\nabla^2 I - \mathbf{w}$ respectively; (b) the distribution of $\nabla I - \mathbf{v}$ (in red), along with a Gaussian fit (in green, $MSE = 4.3085$), and a Laplacian fit (in blue, $MSE = 2.6052$); (b) the distribution of $\nabla^2 I - \mathbf{w}$ (in red), along with a Gaussian fit (in green, $MSE = 6.5732$), and a Laplacian fit (in blue, $MSE = 3.5446$).

Directly minimizing functional (2.2) yields many solutions. If u^* is a minimizer, then $u^* + c$ is also a minimizer, where c is a constant. To avoid this drawback, we add a data-fidelity term which requires that the fused image should be close to a predefined image u_0, which comes from the source images. A quite natural choice for u_0 is $u_0 = \frac{1}{m}\sum_{i=1}^m f_i$. However, this choice may cause contrast loss. For example, if the contrast of the source images are quite different, then by taking average, u_0 probably has low contrast. To tackle this problem, we introduce an index called average contrast (AC) to quantify the contrast of an image I as:

$$AC(I) = \frac{\int_\Omega \int_\Omega |I(x) - I(y)|dxdy}{|\Omega|^2}.$$

Then we choose u_0 to be the source image with highest AC if the difference of ACs for source images are larger than some threshhold, otherwise, we choose the average.

Overall, our resulted model is:

$$\min_u \int_\Omega |\nabla u - s\mathbf{v}| + \alpha|\nabla^2 u - s\mathbf{w}|dx + \frac{\lambda}{2}\int_\Omega |u - u_0|^2 dx, \tag{2.3}$$

where $\alpha \geq 0, \lambda > 0, s \geq 1$. When $s = 1$, the model is for image fusion. When $s > 1$, the magnitude of target first-order and second-order features $\mathbf{v}, \mathbf{w}$ is enlarged, the model can be used for both image fusion and enhancement.

Comparing with the first order method (1), the main theoretical merit of our model is that the existence and uniqueness of minimizer can be proved in some suitable space. Moreover, our general framework can be readily extended. For instance, other second order methods such as TGV^2 [5–7] can also be considered similarly. We leave this as future work.

2.3. Mathematical analysis

In this section, we will give some mathematical analysis of the proposed model (2.3). In the continuous setting, the gradient operator ∇, ∇^2 should be understood as weak derivative in the sense of Radon measure as D, D^2 in $BV(\Omega)$ and $BV^2(\Omega)$ [16, 18, 21]. We remark that $BV(\Omega)$ is a suitable functional space for images as it allows sharp edges. In contrast, at a first glance, $BV^2(\Omega)$ seems a less suitable functional space for describing images since the embedding from $BV^2(\Omega)$ into $W^{1,1}(\Omega)$ is compact. However, as a functional in $BV(\Omega)$, each image can be approximated by smooth functions [18, Theorem 2, page 172]. Moreover, the gap between $BV(\Omega)$ and $W^{1,1}(\Omega)$ will be significantly reduced/removed in the discrete situation where each image has first-order and second-order gradient information defined by finite difference schemes. Indeed, in the numerical aspect, the variational methods involving $BV^2(\Omega)$ have shown to be very effective in image restoration which can well preserve edges while smoothing [13, 21, 28, 31]. Hence, we follow to assume that images belong to $BV^2(\Omega)$ in this paper and leave the extension as future work.

Generally, we let $\Omega \subset \mathbb{R}^n$ be a bounded, open and convex domain with uniform C^1-boundary. Let $V = \mathbb{R}^n$ or $V = \mathbb{R}^{n\times n}$. For $k = 1, 2$, $C_c^k(\Omega, V)$ denotes the set of vector valued, matrix valued, respectively, $k-$times continuously differentiable functions with compact support in Ω and range in V. We denote by

$$E^k(V) := \Big\{\phi \in C_c^k(\Omega, V) : |\phi| \leq 1 \text{ for all } x \in \Omega\Big\}.$$

Here $|\cdot|$ denote Euclidean norm for vector and Frobenius norm for matrix.

For mathematical analysis, we firstly assume that $\mathbf{v}$ and $\mathbf{w}$ are finite vector-valued and matrix-valued Radon measures, i.e.,

$$\int_\Omega |\mathbf{v}| := \sup\left\{\int_\Omega \mathbf{v}\cdot\phi : \phi \in E^\infty(\mathbb{R}^n)\right\} < +\infty, \tag{2.4}$$

$$\int_\Omega |\mathbf{w}| := \sup\left\{\int_\Omega \mathbf{w}\cdot\varphi : \varphi \in E^\infty(\mathbb{R}^{n\times n})\right\} < +\infty. \tag{2.5}$$

In the case of $\alpha = 0$, we rewrite the proposed model (2.3) as:

$$\inf_{u\in BV(\Omega)\cap L^2(\Omega)} E_1(u) = \int_\Omega |Du - s\mathbf{v}| + \frac{\lambda}{2}\int_\Omega |u-u_0|^2 dx. \tag{2.6}$$

In the case of $\alpha > 0$, the model is:

$$\inf_{u\in BV^2(\Omega)\cap L^2(\Omega)} E_2(u) = \int_\Omega |Du - s\mathbf{v}| + \alpha\int_\Omega |D^2u - s\mathbf{w}| + \frac{\lambda}{2}\int_\Omega |u-u_0|^2 dx. \tag{2.7}$$

Remark that

$$\int_\Omega |Du - \mathbf{v}| := \sup\left\{\int_\Omega u \ \mathrm{div}\phi + \mathbf{v}\cdot\phi dx : \phi \in E^\infty(\mathbb{R}^n)\right\}, \tag{2.8}$$

$$\int_\Omega |D^2u - \mathbf{w}| := \sup\left\{\int_\Omega u \ \mathrm{div}^2\varphi + \mathbf{w}\cdot\varphi dx : \varphi \in E^\infty(\mathbb{R}^{n\times n})\right\}. \tag{2.9}$$

Then we can deduce the following inequalities immediately:

$$\int_\Omega |Du| \le \int_\Omega |Du - \mathbf{v}| + \int_\Omega |\mathbf{v}|, \tag{2.10}$$

$$\int_\Omega |D^2u| \le \int_\Omega |D^2u - \mathbf{w}| + \int_\Omega |\mathbf{w}|. \tag{2.11}$$

Then we are ready to prove the existence of a unique minimizer for problem (2.6) and (2.7) respectively.

Theorem 2.1. *Assume that $u_0 \in BV(\Omega)\cap L^2(\Omega), i = 1, \cdots, m$, $\mathbf{v}$ is finite vector-valued Radon measure satisfying* (2.4)*, then the minimization problem* (2.6) *has a unique minimizer $u^* \in BV(\Omega)\cap L^2(\Omega)$.*

Theorem 2.2. *Assume that $u_0 \in BV^2(\Omega)\cap L^2(\Omega), i = 1, \cdots, m, \mathbf{v}$ and $\mathbf{w}$ are finite Radon measures satisfying* (2.4) *and* (2.5)*, $\alpha > 0$, then the minimization problem* (2.7) *has a unique minimizer $u \in BV^2(\Omega)\cap L^2(\Omega)$.*

Since the proof of Theorem 2.1 and Theorem 2.2 are quite similar, we only prove Theorem 2.2 in the following.

Proof. The infimum of energy E_2 must be finite since $E_2(0) < \infty$. Let $\{u_k\}_{k=1}^\infty \in BV^2(\Omega)\cap L^2(\Omega)$ be a minimizing sequence of E_2. Then there exists a constant $C > 0$, such that

$$E_2(u^k) := \int_\Omega |Du^k - \mathbf{v}| + \alpha\int_\Omega |D^2u^k - \mathbf{w}| + \frac{\lambda}{2}\int_\Omega |u^k - u_0|^2 dx \le C. \tag{2.12}$$

Together with inequality (2.10) and (2.11) we deduce that:

$$\int_\Omega |Du^k| \le C, \quad \int_\Omega |D^2u^k| \le C, \tag{2.13}$$

and
$$\int_\Omega |u^k - u_0|^2 dx \le C. \tag{2.14}$$
By using triangle inequalities of L^2 norm on (2.14) and the embedding $L^2(\Omega) \subset L^1(\Omega)$, we can deduce from (2.14) that
$$\int_\Omega |u^k|^2 dx \le C, \quad \int_\Omega |u^k| dx \le C. \tag{2.15}$$
Using the definition of BV^2 norm, we get from (2.13) and (2.15) that $\{u^k\}_{k=1}^\infty$ is bounded in $BV^2(\Omega) \cap L^2(\Omega)$. By the compactness of embedding $BV^2(\Omega) \subset W^{1,1}(\Omega)$ and the reflexive property of $L^2(\Omega)$, we know that there exists a subsequence (again denoted by $\{u^k\}_{k=1}^\infty$) and a function $u^* \in W^{1,1}(\Omega) \cap L^2(\Omega)$ such that
$$\begin{aligned} &u^k \to u^* \;\; \text{strongly in } W^{1,1}(\Omega), \\ &u^k \rightharpoonup u^* \;\; \text{weakly in } L^2(\Omega). \end{aligned}$$
Then by a similar deduction of the lower semi-continuity of bounded variation and second bounded variation, from (2.8) and (2.9) we get that
$$\int_\Omega |Du^* - \mathbf{v}| \le \liminf_{k\to\infty} \int_\Omega |Du^k - \mathbf{v}|, \tag{2.16}$$
$$\int_\Omega |D^2u^* - \mathbf{w}| \le \liminf_{k\to\infty} \int_\Omega |D^2u^k - \mathbf{w}|, \tag{2.17}$$
which in particular implies that $u^* \in BV^2(\Omega)$. Meanwhile by the lower semi-continuity of L^2 norm,
$$\int_\Omega |u^* - u_0|^2 dx \le \liminf_{k\to\infty} \int_\Omega |u^k - u_0|^2 dx, \quad i = 1, \cdots, m. \tag{2.18}$$
Combining inequalities (2.16)-(2.18), on a suitable subsequence, we have established that
$$E_2(u^*) \le \liminf_{k\to\infty} E_2(u^k)$$
and hence u^* is a minimizer.

The uniqueness follows immediately from the strict convexity of E_2. $\square$

3. Numerical Scheme

In this section, we derive the numerical algorithm using alternating split Bregman method which converges and then give the difference scheme in detail.

3.1. The algorithm

In this section, we introduce an efficient algorithm to solve the proposed model based on the popular alternating split Bregman method [11, 20]. Remark that the primal-dual method [14] can also be used to drive an efficient algorithm following [5].

Firstly we give a brief introduction on the split Bregman method. Assume that H and $|\Phi|$

are convex functionals. Let us consider the problem:

$$\min_{u,d} |d| + H(u) \quad s.t. \;\; \Phi(u) = d.$$

Define $F(u,d) = |d| + H(u) + \frac{\mu}{2} \left\| \Phi(u) - d + b^k \right\|_2^2$. The alternating split Bregman algorithm for this problem is given by the following iteration scheme:

$$\begin{aligned} u^{k+1} &= \min_u F(u, d^k), \\ d^{k+1} &= \min_d F(u^{k+1}, d), \\ b^{k+1} &= b^k + \Phi(u^{k+1}) - d^{k+1}. \end{aligned}$$

It has been proved that the alternating split Bregman algorithm converges under some conditions [42].

Without loss of generality, we restrict our analysis on $s = 1$. Let us recall the proposed model:

$$\min_u \int_\Omega |\nabla u - \mathbf{v}| + \alpha |\nabla^2 u - \mathbf{w}| dx + \frac{\lambda}{2} \int_\Omega |u - u_0|^2 dx. \tag{3.1}$$

where all variables are understood as matrices in finite dimension Euclidean space. To implement the alternating split Bregman method on the proposed model (2.3), we first add two auxiliary variables $\mathbf{d}_1, \mathbf{d}_2$ and rewrite (2.3) as:

$$\begin{aligned} \min_{\mathbf{u},\mathbf{d}_1,\mathbf{d}_2} E(\mathbf{u}, \mathbf{v}, \mathbf{w}) &:= \int_\Omega |\mathbf{d}_1| dx + \alpha \int_\Omega |\mathbf{d}_2| dx + \frac{\lambda}{2} \int_\Omega |u - u_0|^2 dx, \\ s.t. \;\; \nabla u - \mathbf{v} &= \mathbf{d}_1, \quad \nabla^2 u - \mathbf{w} = \mathbf{d}_2. \end{aligned}$$

Then using the alternating split Bregman technique on the constraints, we get the iteration scheme:

$$u^{k+1} = \min_u E_r\left(u, \mathbf{d}_1^k, \mathbf{d}_2^k\right), \tag{3.2}$$
$$\left(\mathbf{d}_1^{k+1}, \mathbf{d}_2^{k+1}\right) = \min_{\mathbf{d}_1,\mathbf{d}_2} E_r\left(u^{k+1}, \mathbf{d}_1, \mathbf{d}_2\right), \tag{3.3}$$
$$\mathbf{b}_1^{k+1} = \mathbf{b}_1^k + \nabla u^{k+1} - \mathbf{v} - \mathbf{d}_1^{k+1}, \tag{3.4}$$
$$\mathbf{b}_2^{k+1} = \mathbf{b}_2^k + \nabla^2 u^{k+1} - \mathbf{w} - \mathbf{d}_2^{k+1}, \tag{3.5}$$

where

$$E_r(u, \mathbf{d}_1, \mathbf{d}_2) := \left\{ \begin{array}{l} \int_\Omega |\mathbf{d}_1| dx + \alpha \int_\Omega |\mathbf{d}_2| dx + \frac{\lambda}{2} \int_\Omega |u - u_0|^2 dx \\ + \frac{\mu}{2} \int_\Omega |\nabla u - \mathbf{v} - \mathbf{d}_1 + \mathbf{b}_1^k|^2 dx \\ + \frac{\mu\alpha}{2} \int_\Omega |\nabla^2 u - \mathbf{w} - \mathbf{d}_2 + \mathbf{b}_2^k|^2 dx \end{array} \right\}. \tag{3.6}$$

Here μ is a positive parameter which corresponds to the constraints in (3.1). In the following, we derive the formulas for $u, \mathbf{d}_1$ and $\mathbf{d}_2$ respectively from (3.3) and (3.6) with alternating minimization method.

3.1.1. Solving u

Fixing $\mathbf{d}_1$ and $\mathbf{d}_2$, the subproblem for u is equivalent to:

$$\min_u \left\{ \begin{array}{l} \frac{\lambda}{2} \int_\Omega |u - u_0|^2 dx + \frac{\mu}{2} \int_\Omega |\nabla u - \mathbf{v} - \mathbf{d}_1 + \mathbf{b}_1^k|^2 dx \\ + \frac{\mu\alpha}{2} \int_\Omega |\nabla^2 u - \mathbf{w} - \mathbf{d}_2 + \mathbf{b}_2^k|^2 dx \end{array} \right\}.$$

The corresponding Euler-Lagrangian equation about u is:

$$\lambda(u-u_0)+\mu\nabla^T(\nabla u-\mathbf{v}-\mathbf{d}_1+\mathbf{b}_1^k)+\mu\alpha(\nabla^2)^T(\nabla^2 u-\mathbf{w}-\mathbf{d}_2+\mathbf{b}_2^k)=0,$$

where ∇^T denotes the conjugate operator of ∇ and $(\nabla^2)^T$ denotes the conjugate operator of ∇^2. That is:

$$\left(\gamma+\nabla^T\nabla+\alpha(\nabla^2)^T\nabla^2\right)u=\gamma u_0+\nabla^T(\mathbf{v}+\mathbf{d}_1-\mathbf{b}_1^k)+\alpha(\nabla^2)^T\left(\mathbf{w}+\mathbf{d}_2-\mathbf{b}_2^k\right).$$

with $\gamma=\frac{\lambda}{\mu}$. It can be abbreviated as:

$$Ru=rhs,$$

where

$$R=\gamma+\nabla^T\nabla+\alpha(\nabla^2)^T\nabla^2 \tag{3.7}$$

$$rhs=\gamma u_0+\nabla^T(\mathbf{v}+\mathbf{d}_1-\mathbf{b}_1^k)+\alpha(\nabla^2)^T\left(\mathbf{w}+\mathbf{d}_2-\mathbf{b}_2^k\right). \tag{3.8}$$

Since R can be regarded as convolution operator, under the assumption of periodic boundary condition, we can solve the system efficiently by

$$u=\mathfrak{F}^{-1}\left(\frac{\mathfrak{F}\,(rhs)}{\mathfrak{F}\,(R)}\right) \tag{3.9}$$

where $\mathfrak{F}$ denotes the fast Fourier transform (FFT) and $\mathfrak{F}^{-1}$ denotes the inverse fast Fourier transform (IFFT). See section 3.2 for more details.

3.1.2. Solving $\mathbf{d}_1$ and $\mathbf{d}_2$

Fixing u, since the subproblems for $\mathbf{d}_1$ and $\mathbf{d}_2$ are separable, we can solve them one by one. The subproblem for $\mathbf{d}_1$ is equivalent to:

$$\min_{\mathbf{d}_1}\left\{\int_\Omega|\mathbf{d}_1|dx+\frac{\mu}{2}\int_\Omega|\nabla u-\mathbf{v}-\mathbf{d}_1+\mathbf{b}_1^k|^2dx\right\}.$$

A simple calculation gives the closed-form solution of this problem:

$$\mathbf{d}_1=\max\left\{\left|\mathbf{b}_1^k+\nabla u-\mathbf{v}\right|-1/\mu,0\right\}\frac{\mathbf{b}_1^k+\nabla u-\mathbf{v}}{\left|\mathbf{b}_1^k+\nabla u-\mathbf{v}\right|}. \tag{3.10}$$

The subproblem for $\mathbf{d}_2$ is equivalent to:

$$\min_{\mathbf{d}_2}\left\{\int_\Omega|\mathbf{d}_2|dx+\frac{\mu}{2}\int_\Omega|\nabla^2 u-\mathbf{w}-\mathbf{d}_2+\mathbf{b}_2^k|^2dx\right\}.$$

Similar calculation gives the following closed-form solution:

$$\mathbf{d}_2=\max\left\{\left|\mathbf{b}_2^k+\nabla^2 u-\mathbf{w}\right|-1/\mu,0\right\}\frac{\mathbf{b}_2^k+\nabla^2 u-\mathbf{w}}{\left|\mathbf{b}_2^k+\nabla^2 u-\mathbf{w}\right|}. \tag{3.11}$$

Finally, we summarize the algorithm as in Program 1.

Program 1

- Initialization: $u^0 = \mathbf{0}, \mathbf{b}_1^0 = \mathbf{0}, \mathbf{b}_2^0 = \mathbf{0}, \mathbf{d}_1^0 = \mathbf{0}, \mathbf{d}_2^0 = \mathbf{0}$.
- For $k = 0, 1, 2, \ldots$, repeat until a stopping criterion is reached

$$u^{k+1} = \mathfrak{F}^{-1}\left(\frac{\mathfrak{F}(rhs)}{\mathfrak{F}(R)}\right), \text{see (3.7), (3.8) for the formula of } R \text{ and } rhs,$$
$$\mathbf{d}_1^{k+1} = \max\left\{\left|\mathbf{b}_1^k + \nabla u^{k+1} - \mathbf{v}\right| - 1/\mu, 0\right\} \frac{\mathbf{b}_1^k + \nabla u^{k+1} - \mathbf{v}}{\left|\mathbf{b}_1^k + \nabla u^{k+1} - \mathbf{v}\right|},$$
$$\mathbf{d}_2^{k+1} = \max\left\{\left|\mathbf{b}_2^k + \nabla^2 u^{k+1} - \mathbf{w}\right| - 1/\mu, 0\right\} \frac{\mathbf{b}_2^k + \nabla^2 u^{k+1} - \mathbf{w}}{\left|\mathbf{b}_2^k + \nabla^2 u^{k+1} - \mathbf{w}\right|},$$
$$\mathbf{b}_1^{k+1} = \mathbf{b}_1^k + \nabla u^{k+1} - \mathbf{v} - \mathbf{d}_1^{k+1},$$
$$\mathbf{b}_2^{k+1} = \mathbf{b}_2^k + \nabla^2 u^{k+1} - \mathbf{w} - \mathbf{d}_2^{k+1}.$$

- Output: u^{k+1}.

Theorem 3.1. *The sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by the alternating split Bregman algorithm in Program* 1 *converges to the minimizer of the proposed energy in* (3.1)*.*

Proof. With a standard argument similar to the proof of Theorem 2.2, we can prove that in discrete setting the functional in (3.1) is proper, convex, restrictive and lower semi-continuous, and thus the uniqueness of minimizer is guaranteed. According to Theorem 5 and Corollary 1 in [42], the convergence result holds. □

3.2. Difference scheme

In discrete setting, we will utilize a two-dimensional regular Cartesian grid of size $N \times N$:

$$\Omega = \{(i,j) | i = 1, \cdots, N, j = 1, \cdots, N\}$$

where (i,j) denotes a pixel of the image. In this paper, we utilize both the two-direction difference scheme and the four-direction difference scheme for first-order gradient operator. Let us define the following difference operators at $x, y, 45^\circ$ and 135° directions with periodic boundary condition:

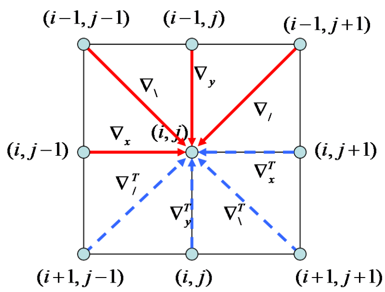
$$\nabla_x u(i,j) = \begin{cases} u(i,j) - u(i,j-1), & i = 1, \cdots, N, j = 2, \cdots, N \\ u(i,1) - u(i,N), & i = 1, \cdots, N, j = 1. \end{cases}$$
$$\nabla_y u(i,j) = \begin{cases} u(i,j) - u(i-1,j), & i = 2, \cdots, N, j = 1, \cdots, N \\ u(1,j) - u(N,j), & i = 1, j = 1, \cdots, N. \end{cases}$$

$$\nabla_{/}u(i,j)=\begin{cases}[u(i,j)-u(i-1,j+1)]/\sqrt{2}, & i=2,\cdots,N, j=1,\cdots,N-1\\ [u(1,j)-u(N,j+1)]/\sqrt{2}, & i=1, j=1,\cdots,N-1\\ [u(i,N)-u(i-1,1)]/\sqrt{2}, & i=2,\cdots,N, j=N\\ [u(1,N)-u(N,1)]/\sqrt{2}, & i=1, j=N.\end{cases}$$

$$\nabla_{\backslash}u(i,j)=\begin{cases}[u(i,j)-u(i-1,j-1)]/\sqrt{2}, & i=2,\cdots,N, j=2,\cdots,N\\ [u(1,j)-u(N,j-1)]/\sqrt{2}, & i=1, j=2,\cdots,N\\ [u(i,1)-u(i-1,N)]/\sqrt{2}, & i=2,\cdots,N, j=1\\ [u(1,1)-u(N,N)]/\sqrt{2}, & i=1, j=1.\end{cases}$$

Periodic boundary condition is chosen to enable the use of FFT. The corresponding conjugate of the above four operators are then simply the negative finite differences. See Fig. 3.1 for an intuitive description of these operators.

Remark that ∇_x, ∇_y are backward difference scheme widely used in image community to discrete the gradient operator in image restoration problems [12]. While the last two operators $\nabla_{/}, \nabla_{\backslash}$ are rarely used. Actually, we have tested that in the well know total variation denoising model [40], four-direction difference scheme has similar performance as the common used two-direction difference scheme. However, as will be shown in Section 4,the four-direction scheme seems superior to two-direction scheme in image fusion problem. One possible reason is that the role of gradient term in image fusion is something different from that in image denoising. Indeed, in image denoising, $\int_\Omega |\nabla u| dx$ is a regularization term to smooth the solution image, and both versions of difference scheme approximate to the same total variation value. However, in image fusion, $\int_\Omega |\nabla u - \mathbf{v}| dx$ is a fidelity term which requires that ∇u matches with $\mathbf{v}$, where more directions on difference provide extra information as our feature selection procedure to build $\mathbf{v}, \mathbf{w}$ (see Section 2.1) is nonlinear.



(a)

Fig. 3.1. The first-order difference scheme. The difference operators at $x, y, 45^\circ$ and 135° directions, $\nabla_x, \nabla_y, \nabla_{/}$ and $\nabla_{\backslash}$, are vectors in sold line in red; the corresponding conjugate operators, $\nabla_x^T, \nabla_y^T, \nabla_{/}^T$, and $\nabla_{\backslash}^T$, are vectors in dashed line in blue.

The second-order difference operators can be obtained though the composition of first-order difference operators. For example, in the two-direction case, let $\nabla=(\nabla_x,\nabla_y)$, correspondingly $\nabla^T=(\nabla_x^T,\nabla_y^T)$, then the second-order difference operators are

$$\nabla^2=\begin{pmatrix}\nabla_x\nabla_x, \nabla_x\nabla_y,\\ \nabla_y\nabla_x, \nabla_y\nabla_y\end{pmatrix}, \quad (\nabla^2)^T=\begin{pmatrix}\nabla_x^T\nabla_x^T, \nabla_y^T\nabla_x^T,\\ \nabla_x^T\nabla_y^T, \nabla_y^T\nabla_y^T\end{pmatrix}.$$

There are three individual operators in ∇^2 and $\left(\nabla^2\right)^T$ since the derivatives can change turn, i.e., $\nabla_x\nabla_y = \nabla_y\nabla_x$ and $\nabla_y^T\nabla_x^T = \nabla_x^T\nabla_y^T$. Similarly, if four-direction difference scheme is used for first-order gradient operator, i.e., $\nabla = (\nabla_x, \nabla_y, \nabla_/, \nabla_\backslash)$, correspondingly $\nabla^T = (\nabla_x^T, \nabla_y^T, \nabla_/^T, \nabla_\backslash^T)$, then the second-order operators are

$$\nabla^2 = \begin{pmatrix} \nabla_x\nabla_x, \nabla_x\nabla_y, \nabla_x\nabla_/, \nabla_x\nabla_\backslash, \\ \nabla_y\nabla_x, \nabla_y\nabla_y, \nabla_y\nabla_/, \nabla_y\nabla_\backslash, \\ \nabla_/\nabla_x, \nabla_/\nabla_y, \nabla_/\nabla_/, \nabla_/\nabla_\backslash, \\ \nabla_\backslash\nabla_x, \nabla_\backslash\nabla_y, \nabla_\backslash\nabla_/, \nabla_\backslash\nabla_\backslash \end{pmatrix}, \quad \left(\nabla^2\right)^T = \begin{pmatrix} \nabla_x^T\nabla_x^T, \nabla_y^T\nabla_x^T, \nabla_/^T\nabla_x^T, \nabla_\backslash^T\nabla_x^T, \\ \nabla_x^T\nabla_y^T, \nabla_y^T\nabla_y^T, \nabla_/^T\nabla_y^T, \nabla_\backslash^T\nabla_y^T, \\ \nabla_x^T\nabla_/^T, \nabla_y^T\nabla_/^T, \nabla_/^T\nabla_/^T, \nabla_\backslash^T\nabla_/^T, \\ \nabla_x^T\nabla_\backslash^T, \nabla_y^T\nabla_\backslash^T, \nabla_/\nabla_\backslash^T, \nabla_\backslash^T\nabla_\backslash^T \end{pmatrix}. \tag{3.12}$$

There are ten individual operators in ∇^2 and $\left(\nabla^2\right)^T$ respectively. In numerical implementation, we take use of this symmetric property to reduce the computational time. Remark that we display ∇^2 in matrix form, but we rearrange it to vector form in the model (2.3) and the algorithm for simplicity.

Since we use FFT to solve u in the algorithm, to calculate $\mathfrak{F}(R)$ in Program 1, we need the form of first-order difference operators $\nabla_x, \nabla_y, \nabla_/, \nabla_\backslash$ as convolution kernels. It can be directly obtained from their definitions that the kernels are

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \quad \frac{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}{\sqrt{2}}, \quad \frac{\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}}{\sqrt{2}}.$$

The corresponding higher order operators as convolution kernels can be obtained by convolution of the first-order kernels. For example, the convolution kernel of $\nabla_x\nabla_y$ is

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

4. Numerical Results

In this section, we test our algorithm on several sets of source images. Some of the experimental results are compared with existing popular algorithms including Laplacian pyramid, Gradient pyramid, discrete wavelet transform pyramid, first-order variational model (1.1), methods in [25] and [45]. For simplicity the first five methods are abbreviated as *Laplacian, Gradient, DWT* and *Order1-old* respectively. Remark that for fair comparison, we use the same feature selection rule as described in Section 2.1 in high frequency for the pyramid methods and the proposed method. To choose low frequency features in the pyramid methods, we use the commonly used average method. For variational method *Order1-old*, we follow the setting in [36].

In the proposed algorithm, we test three cases:

Order1-2: first-order method with $\alpha = 0$, $\nabla = (\nabla_x, \nabla_y)$.

Order1-4: first-order method with $\alpha = 0$, $\nabla = (\nabla_x, \nabla_y, \nabla_/, \nabla_\backslash)$.

Order2: second-order method with $\alpha > 0$, $\nabla = (\nabla_x, \nabla_y, \nabla_/, \nabla_\backslash)$ and ∇^2 as in (3.12).

We use the following parameters setting for gray level image fusion in the proposed method: window size $= 5 \times 5$ for feature selection, $\lambda = 0.01, \mu = 0.5$. Specially we set iteration $= 5$ in

Order1-2 and *Order1-4*, while $\alpha = 0.02$, iteration $= 6$ in *Order2*. These parameters are chosen by trail-and-error in order to achieve the optimal results.

Remark that all source images in the following experiments except Cameraman in Fig. 4.1 are downloaded from the website [19]. There also include some introductions of the source images.

All the experiments are performed under Windows 8 and MATLAB R2012a with Intel Core i7-4500 CPU@1.80GHz and 8GB memory. The programming language is MATLAB.

4.1. Test 1

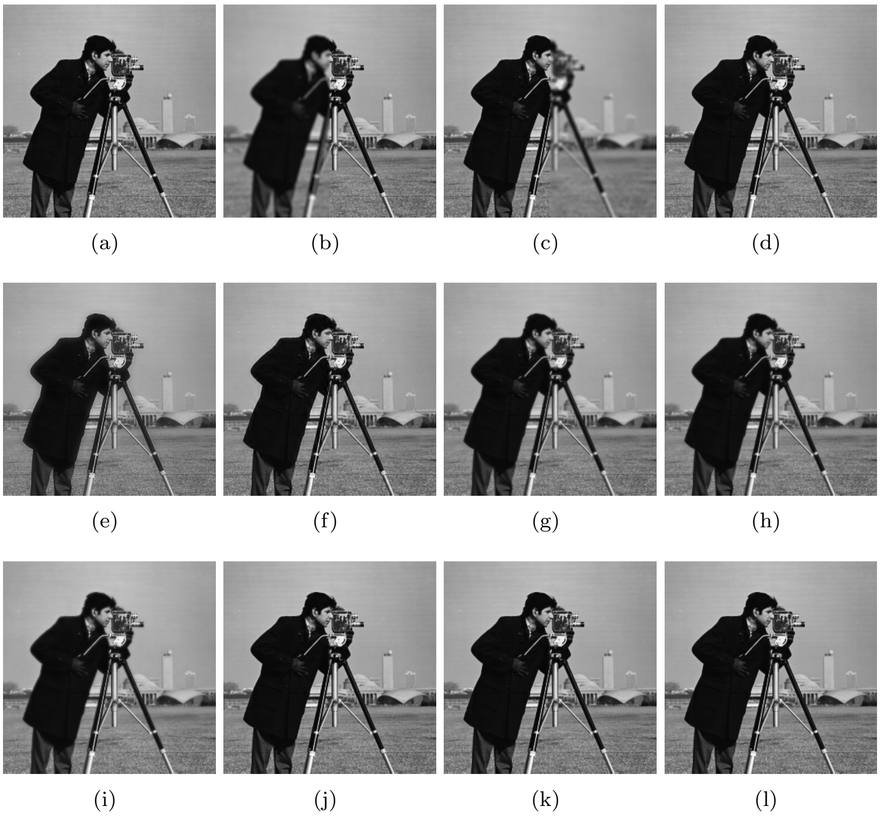


Fig. 4.1. Test on synthetic images. (a)-(c) test images with size 256×256: (a) the ground truth image of Cameraman; (b) and (c) input source images by blurring the left part and right part of (a); (d)-(i) fused images obtained by the existing methods: (d) *Laplacian*; (e) *Gradient*; (f) *DWT*; (g) method in [25]; (h) method in [45]; (g) *Order1-old*; (j)-(l) fused images obtained by the proposed methods: (j) *Order1-2*; (k) *Order1-4*; (l) *Order2*.

Firstly we test synthetic images in Fig. 4.1 by the proposed method and some other methods. Fig. 4.1(a) is the ground truth Cameraman image. Fig. 4.1(b) and Fig. 4.1(c) are input source images to be fused. They are created by blurring (a) in the left part and right part respectively with a Gaussian kernel with mean zero and standard deviation $\sigma = 2$ (kernel size 5×5). The source images are complementary. By careful observation, we find that Fig. 4.1(e), the result of gradient pyramid, has lost some contrast as the whole image is dark. While the edge details of Fig. 4.1(g)-(i), the results of variational model (1.1) and methods in [25] and [45], seem not salient compared with Fig. 4.1(d), Fig. 4.1(f) and Fig. 4.1(j)-(l).

In order to see the difference more clearly, we display the residues of ground truth image with the fused images in Fig. 4.2. It is obvious that Fig. 4.2(b) contains much contrast and edge information, Fig. 4.2(c)-(f) contain some edge information. However the Laplacian pyramid and the proposed method lost very little information in the residues except along the middle line, which is the boundary of true information and damaged information in source images. It suggests that Laplacian pyramid and the proposed method have recovered more salient features such as edges than other methods. Comparing Fig. 4.2(a) and Fig. 4.2(g)-(i), the residues are similar. However, by careful observation we find that Fig. 4.2(a) and Fig. 4.2(g) lost more information along the middle line than Fig. 4.2(h)-(i).

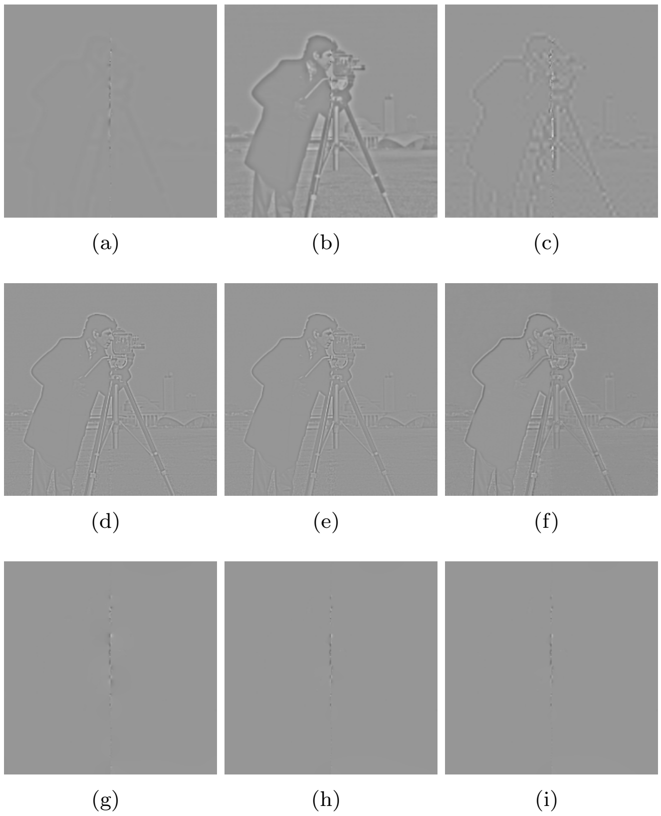


Fig. 4.2. The residues of ground truth image I and the fused images u, display $I - u + 150$. (a)-(f) residues of the exiting methods (a) *Laplacian*; (b) *Gradient*; (c) *DWT*; (d) method in [25]; (e) method in [45]; (f) *Order1-old*; (g)-(i) residues of the proposed methods: (g) *Order1-2*; (h) *Order1-4*; (i) *Order2*.

For quantitative comparison, we calculate image quality measures illustrated in Table 4.1. Since in this test, we have ground truth, so peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are standard performance measures. Higher PSNR is better. For SSIM, the maximum value is 1 and larger is better. We also calculate three popular fusion quality measures requiring no reference image including: MI – which is mutual information [39], $Q_{AB\backslash F}$ – which takes into account the edge strength and orientation preserving values [48], and Q_E –

which takes into account the locations as well as magnitude of distortions [36, 37, 46]. Remark that $MI, Q_{AB\backslash F}, Q_E$ are normalized and ideal values are all 1, and bigger is better.

In each row of Table 4.1, the bold-faced number is the best one and the italic number is the second best. Among all, *Order2* achieves all the best quality measures, and *Order1-4* gains all the second best measures. The pyramid method *Laplacian* and *Order1-2* gain similar performance measures. *Gradient* has the lowest quality measures which coincides with its poor visual quality. The performance of methods in [25], [45] and *Order1-old* are similar which are better than *Gradient* in every quality measure. *DWT* ranks in the middle. Let us pay more attention to measures PSNR and SSIM which are widely used when the reference image is available. *Laplacian* and *Order1-2* have similar PSNR about 46.5dB, which are much higher than *Gradient*, [25], [45], *Order1-old* and *DWT*. As four directions are used in first-order difference scheme, *Order1-4* dramatically improves the PSNR about 1dB. Furthermore, if the second-order derivatives are also considered, *Order2* achieves a PSNR about 0.6 dB higher than *Order1-4*. SSIM measures seem almost consistent with PSNR. *Gradient* and *Order1-old* have the lowest SSIM. [25], [45], *DWT* and *Laplacian* are ranked in the middle. The proposed methods outperform others in SSIM index, among which *Order2* is the best. We remark that SSIM is quite similar for the proposed methods which are very close to the maximum 1.

We also report the computational time (average of ten times running) of each method in Table 4.1. The noniterative methods *Laplacian*, *Gradient* and *DWT* are much faster than others which takes less than 0.04s. The method in [25] is most time consuming (3.4112s) since it process the image pixel by pixel. The iterative method in [45] takes about 1.5 seconds. *Order1-old* and the proposed methods takes less than 1 second. *Order1-old* is slower than our methods since gradient descent method is used in the implementation which needs about one hundred iterations, while our methods take only 5-6 iterations. Among the proposed methods, *Order2* is the most time-consuming and *Order1-2* is the fastest. It is obvious that the computational time is closely related with the number of directions involved in the difference scheme. With more directions, the algorithm is more time-consuming in one iteration.

Table 4.1: Quality measures and computational time for different tested fusion methods corresponding to Fig. 4.1.

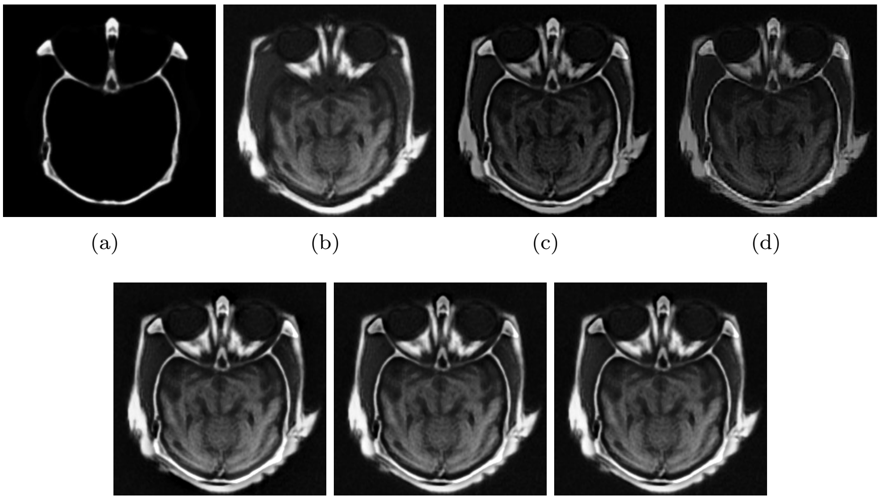
	Laplacian	*Gradient*	*DWT*	[25]	[45]	*Order1-old*	*Order1-2*	*Order1-4*	*Order2*
$PSNR$	46.5143	30.7681	37.2532	32.7390	33.6798	31.3811	46.5163	*47.6896*	**48.2745**
$SSIM$	0.9991	0.9708	0.9922	0.9796	0.9850	0.9678	0.9993	*0.9994*	**0.9995**
MI	0.5838	0.4554	0.5082	0.5181	0.5233	0.4753	0.5968	*0.6141*	**0.6169**
$Q_{AB\backslash F}$	0.7434	0.7026	0.7104	0.7423	0.7455	0.7321	0.7460	*0.7467*	**0.7470**
Q_E	0.7604	0.7329	0.7513	0.7391	0.7496	0.7406	0.7607	*0.7609*	**0.7610**
$Time(s)$	0.0318	0.0123	0.0178	3.4112	1.5326	0.9018	0.1082	0.1843	0.8840

4.2. Test 2

From test 1 we know that except our methods, *Laplacian* and *DWT* perform better than others. Therefore in this test we limit our comparison with these two methods in the fusion of two medical images. Fig. 4.3(a) shows the CT image in which the structure of bone is better visualized, while Fig. 4.3(b) is the MR image in which the pathological soft tissues are better visualized. By fusing them into a single image, one can obtain salient features in both modalities and at the same time display the relative position of soft tissues and bone structure. Fig. 4.3(c)-(g) show the fused results. By careful observation, we find that there are some artifacts in the fused image by *DWT* especially along the bone in the bottom of the image.

Moreover, the results of *Laplacian* and DWT show some extent of contrast loss. While by our methods, the contrast is better preserved in Fig. 4.3(e)-(g). The reason is that the source image Fig. 4.3(b) has much higher AC than Fig. 4.3(a) and thus the former is chosen to be u_0. For further comparison, we report the fusion measures $MI, Q_{AB\backslash F}$ and Q_E. Remark that there is no reference image in this case. Table 4.2 shows that: *Order2* achieve the best performance, *Order1-4* gains the second best performance of MI and Q_E. DWT has the worst performance. *Laplacian* performs better than DWT but worse than the proposed methods.

The computational time of each method is reported in Table 4.2. The noniterative methods *Laplacian* and DWT are much faster than others. Among the proposed methods, similar to Test 1, we find that *Order2* is the most time-consuming and *Order1-2* is the fastest.



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

(e) (f) (g)

Fig. 4.3. Test on medical images. First row: the input source images with size 256 × 256: (a) CT image; (b) MRI image, fused images obtained by (c) *Laplacian*; (d) DWT; Second row: fused image obtained by (e) *Order1-2*; (f) *Order1-4*; (g) *Order2*.

Table 4.2: Quality measure and computational time for different tested fusion methods corresponding to Fig. 4.3.

	Laplacian	DWT	*Order1-2*	*Order1-4*	*Order2*
MI	0.3416	0.3676	0.3731	*0.4421*	**0.4583**
$Q_{AB\backslash F}$	0.7284	0.5156	*0.7941*	0.7930	**0.7954**
Q_E	0.5753	0.4627	0.6938	*0.6981*	**0.6999**
$Time(s)$	0.0139	0.0189	0.1156	0.2017	0.9116

4.3. Test 3

In the following, we display more results on various tasks of image fusion and enhancing to demonstrate the effectiveness of the proposed model. Here we report the results of *Order1-4* and *Order2*.

In Fig. 4.4, we test three pairs of images including multi-focus images, aircraft images and remote sensing images. The first and second columns are the test image pairs. The third and fourth column show the fusion results of the proposed methods *Order1-4* and *Order2*

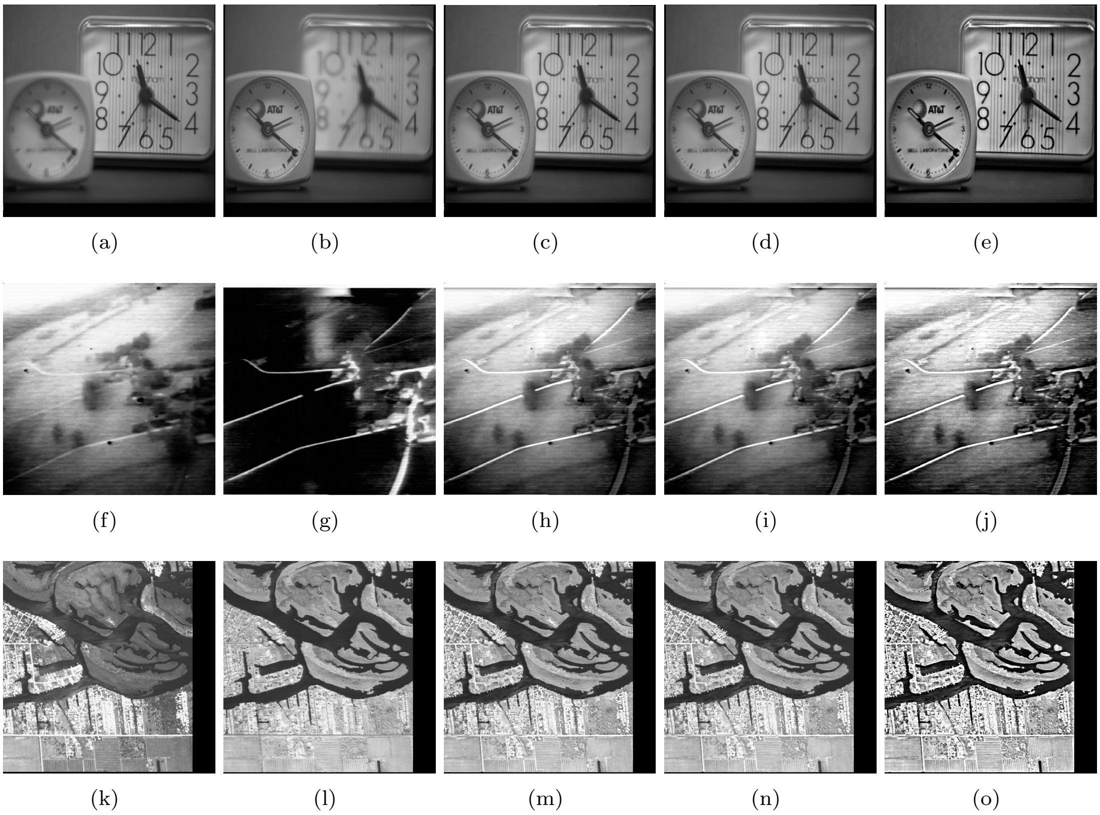
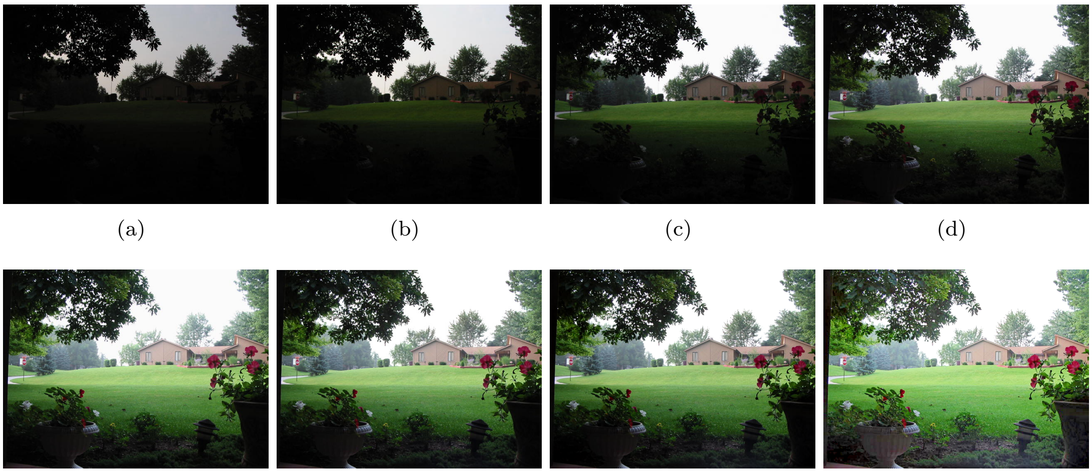


Fig. 4.4. Fusion of more images with the proposed methods. First row is multi-focus image fusion: (a) source image focused in the right clock; (b) source image focused in the left clock; (c) the fused image by *Order1-4*; (d) the fused image by *Order2*; (e) the fused and enhanced image by *Order2*; Second row is aircraft navigation image fusion: (f) LLTV image; (g) FLIR image; (h) the fused image by *Order1-4*; (i) the fused image by *Order2*; (j) the fused and enhanced image by *Order2*; Third row is remote sensing image fusion: (k) and (l) are two bands of source images obtained by multispectral scanner; (m) the fused image by *Order1-4*; (n) the fused image by *Order2*; (o) the fused and enhanced image by *Order2*.

respectively. Since the performance of these two methods are quite similar, we choose to display the enhanced results of *Order2* only in the last column. Let us give more details of Fig. 4.4 in the following. In the first row of Fig. 4.4, fusion of multi-focus images is tested. Fig. 4.4(a) and Fig. 4.4 (b) show an image pair focused in right clock and left clock respectively. In the fused image Fig. 4.4(c)-(d), which display the fusion results of the proposed methods *Order1-4* and *Order2* respectively, both of the clocks are in focus and clear. In the enhanced version Fig. 4.4(e) by method *Order2*, the numbers in the clocks are more highlighted. The second row shows fusion of aircraft navigation images. To allow helicopter pilots navigate under poor visibility conditions (such as fog or heavy rain) helicopters are equipped with several imaging sensors, which can be viewed by the pilot in a helmet mounted display. For example, Fig. 4.4(f) is the source image captured using a low-light-television (LLTV) sensor and Fig. 4.4(g) is the source image obtained using a forward-looking-infrared (FLIR) sensor. Note that LLTV sensor provides the surface information of the ground, building and vegetation details around. While FLIR sensor provides the information of road network details accurately. In the fused image

Fig. 4.4(h)-(i), all the salient features are clear. In Fig. 4.4(j), the details of the ground and the roads are enhanced and better visualized. The third row shows the fusion of two bands of multispectral scanner images in Fig. 4.4(k)-(l). Band 1 penetrates water, which is useful for mapping along costal areas and distinguishing soft-vegetation and forest type. Band 2 is good at detecting green vegetation water-land interface [36]. The fused image Fig. 4.4(m)-(n) combines the useful salient features in both bands and obviously enhanced in Fig. 4.4(o).

4.4. Test 4



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

(e) (f) (g) (h)

Fig. 4.5. Multi-exposure images fusion. (a)-(e) Five source images f_1, f_2, f_3, f_4 and f_5; (f) the fused image by the proposed method *Order1-4*; (g) the fused image by the proposed method *Order2*; (h) the enhanced image of fused result (g).

This example is multi-exposure images fusion. It is sometimes impossible to obtain a single image of a scene where all areas appear well-exposed. Some scene areas may appear under-exposed or over-exposed in a single shot. Each local area in a scene may require a different shutter speed to best capture its details [19]. By fusion, we combine the five images of a scene captured at different shutter speeds into a single image where all scene areas appear well exposed, see Fig. 4.5(f)-(g) for the fusion results of the proposed method *Order1-4* and *Order2* respectively. The results of these two methods are quite similar. In this example, as shown in Fig. 4.5(e), f_5 has the largest average contrast. So we set $u_0 = f_5$ in the algorithm. For better visualization, we apply the variational model for retinex [33] on Fig. 4.5(g) to enhance the image (since Fig. 4.5(f)-(g) are quite similar, we take the latter as an example). The enhanced result is shown in Fig. 4.5(h).

Two parts around the house are enlarged in Fig. 4.6. The first and second rows show the enlarged regions of the source images. The last row show the enlarged regions of the fused images by the proposed method *Order1-4* and *Order2* respectively. The results are quite similar. The leaves are better visualized in Fig. 4.6(k)-(l), since they have more details than Fig. 4.6(e) and they have no artifacts as in Fig. 4.6(a)-(d) around the leaves. In Fig. 4.6(m)-(n), the trees behind the house are better visualized than in Fig. 4.6(j) since the formers include most salient details in source images.

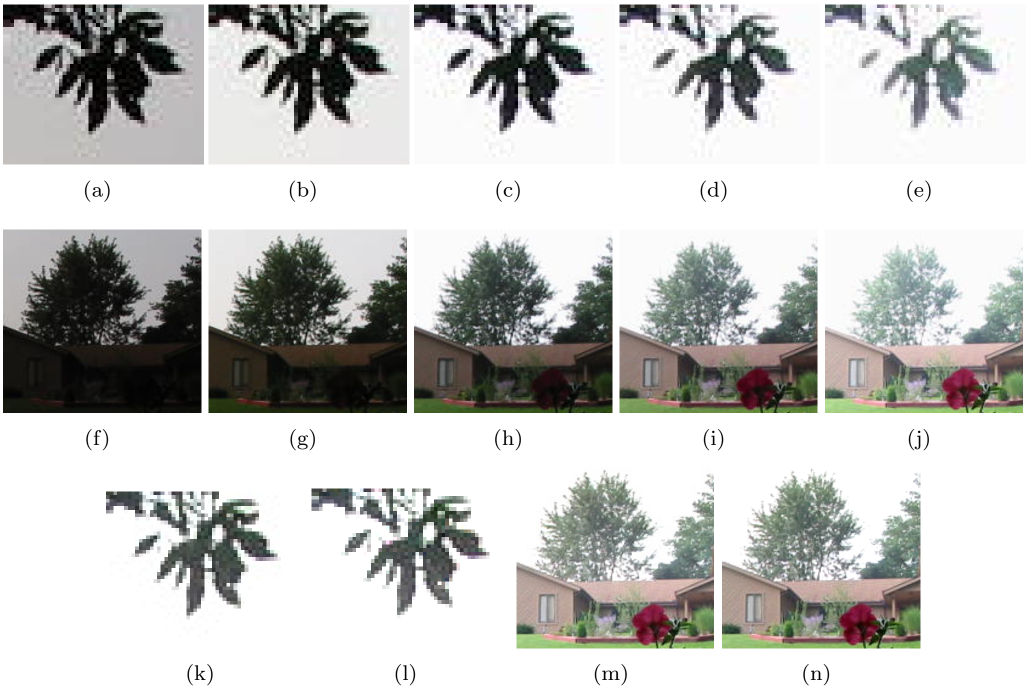


Fig. 4.6. Two parts of the source images and fused result are zoomed for comparison. (a)-(e) and (f)-(j) are parts of source images in Fig. 4.5(a)-(e); (k) and (l) are parts of fused image in Fig. 4.5(f) by the proposed method *Order1-4*, respectively; (m) and (n) are parts of fused image in Fig. 4.5(g) by the proposed method *Order2*, respectively.

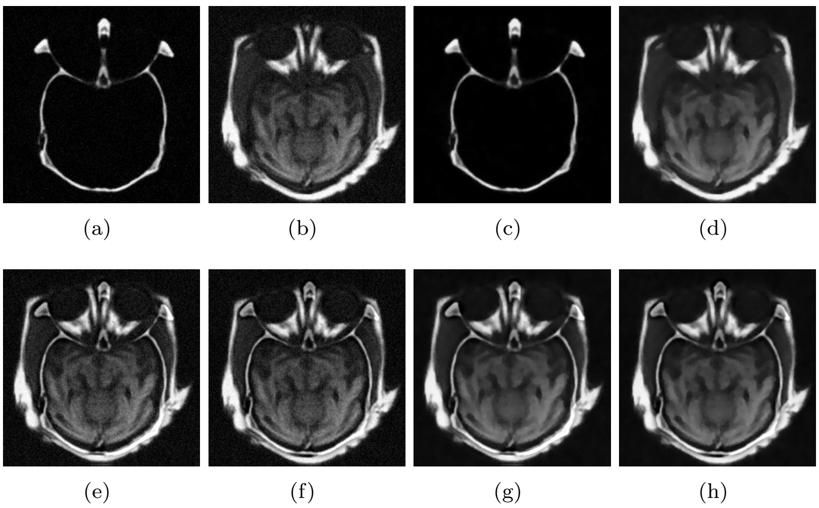
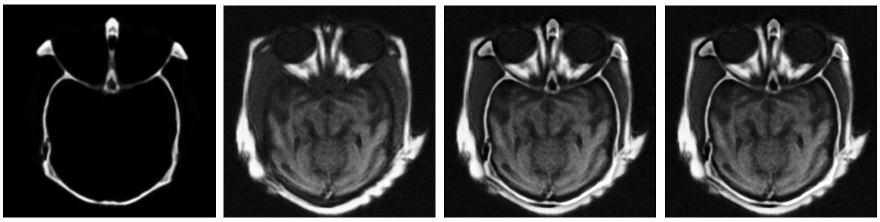


Fig. 4.7. Medical image fusion with noise. (a)-(b) The noisy source images are noisy versions of Fig. 4.3(a)-(b) which are contaminated by Gaussian noise with standard deviation 10 and mean zero; (c)-(d) the denoised source images by solving total variation denoising model [40] with split Bregman method; (e)-(f) the fusion results of the proposed methods *Order1-4* and *Order2* on noisy source images respectively; (g)-(h) the fusion results of the proposed methods *Order1-4* and *Order2* on denoised source images (c) and (d), respectively.



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

Fig. 4.8. Medical image fusion with miss match. The source images (a) is the as Fig. 4.3(a); The source image (b) is obtained by shifting Fig. 4.3(b) 2 pixels along x axis and y axis respectively; (c)-(d) the fusion result of the proposed methods *Order1-4* and *Order2* respectively.

4.5. Test 5

In the last example, we test the robustness of the proposed method for images with small noise and miss match. In Fig. 4.6, we add Gaussian noise with standard deviation 10 on the source images which are tested in Fig. 4.3. Without preprocessing, the results of the proposed methods *Order1-4* and *Order2* are somewhat noisy, see Fig. 4.6(c)-(d). However, after preprocessed by total variation denoising method in [40], the results are very good as shown in Fig. 4.6(e)-(f). In Fig. 4.7, we test the robustness of the proposed methods when the source images are slightly miss match. The two source images in Fig. 4.7(a)-(b) are miss match by shifting 2 pixels along each axis. From the fusion results in Fig 4.7(c)-(d), we find that the miss match affects the fusion results slightly.

5. Conclusions

We have presented in this paper a new variational approach to image fusion. The contribution of our paper is clear. Indeed, both the first-order and second-order gradient information at different directions have been considered as features in our framework. And a new feature selection rule has been built. Mathematically we established the existence of unique minimizer for the proposed model. The proposed four-direction difference scheme for gradient operator seems promising in image fusion. More important, our variational framework (2.3) can be readily extended in various aspects. In the future work, we will study high order regularization technique TGV^2 [5] in the fusion process which should be more suitable for images with edges and consider to handle the contrast loss problem in a variational framework such as in [36] rather than using a preprocessing method. We will also generalize the proposed model to other image fusion applications such as pan-sharpening.

Acknowledgments. This work is supported by the 973 Program (2011CB707104), the Science and Technology Commission of Shanghai Municipality (STCSM) 13dz2260400, the National Science Foundation of China (Nos. 11001082, 11271049), and RGC 211710, 211911, 12302714 and RFGs of HKBU.

References

[1] A. Ardeshir Goshtasby and S. Nikolov. Image fusion: Advances in the state of the art. *Inform. Fusion*, **8**:2 (2007), 114–118.

[2] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé. A variational model for p+xs image fusion. *Int. J. Comput. Vision*, **69**:1 (2006), 43–58.

[3] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi. Perceptual color correction through variational techniques. *IEEE Trans. Image Process.*, **16**:4 (2007), 1058–1072.

[4] R.S. Blum and Zheng. Liu. *Multi-sensor image fusion and its applications, (special series on Signal Processing and Communications).* Taylor & Francis, CRC Press, 2005.

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

[6] Kristian Bredies, Karl Kunisch, and Tuomo Valkonen. Properties of $l^1 - tgv^2$: The one-dimensional case. *J. Math. Anal. Appl.*, **398**:1 (2013), 438–454.

[7] Kristian Bredies and Tuomo Valkonen. Inverse problems with second-order total generalized variation constraints. *Proceedings of SampTA*, 201, 2011.

[8] Ninon Burgos, Manuel Jorge Cardoso, Marc Modat, Stefano Pedemonte, John Dickson, Anna Barnes, John S Duncan, David Atkinson, Simon R Arridge, Brian F Hutton, et al. Attenuation correction synthesis for hybrid pet-mr scanners. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 147–154. Springer, 2013.

[9] P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Trans. Commun.*, **31**:4 (1983), 532–540.

[10] P.J. Burt and R.J. Kolczynski. Enhanced image capture through fusion. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*, pages 173–182. IEEE, 1993.

[11] J.F. Cai, S. Osher, and Z. Shen. Split bregman methods and frame based image restoration. *Multiscale Model. Simul.*, **8**:2 (2009), 337.

[12] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, **20**:1 (2004), 89–97.

[13] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, **76**:2 (1997), 167–188.

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

[15] Y. Chen and SE Levine. Image recovery via diffusion tensor and time-delay regularization. *J. Vis. Commun. Image. R.*, **13**:1(2002), 156–175.

[16] F. Demengel. Fonctions à hessien borné. *Ann. Inst. Fourier*, **34**:2 (1984), 155–190.

[17] Matthias Joachim Ehrhardt and Simon R Arridge. Vector-valued image processing by parallel level sets. *IEEE Trans. Image Process.*, **23**:1 (2014), 9–18.

[18] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions.* CRC, 1992.

[19] The Image fusion server. http://www.imagefusion.org/.

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

[21] W. Hinterberger and O. Scherzer. Variational methods on the space of functions of bounded hessian for convexification and denoising. *Computing*, **76**:1 (2006), 109–133.

[22] Y. Hu and M. Jacob. Higher degree total variation (hdtv) regularization for image recovery. *IEEE Trans. Image Process.*, **21**:5 (2012), 2559–2571.

[23] S. Kluckner, T. Pock, and H. Bischof. Exploiting redundancy for aerial image fusion using convex optimization. *Pattern Recognition Lecture Notes in Computer Science*, 2010, 303–312.

[24] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *Adv. Neur. In.*, pages 2009, 1033–1041.

[25] M. Kumar and S. Dass. A total variation-based algorithm for pixel-level image fusion. *IEEE Trans. Image Process.*, **18**:9 (2009), 2137–2143.

[26] E. Lallier and M. Farooq. A real time pixel-level based image fusion via adaptive weight averaging. In *Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on*, volume 2, pages WEC3–3. IEEE, 2000.

[27] F. Li, L. Pi, and T. Zeng. Explicit coherence enhancing filter with spatial adaptive elliptical kernel. *IEEE Signal. Proc. Let.*, **19**:9 (2012), 555–558.

[28] F. Li, C. Shen, J. Fan, and C. Shen. Image restoration combining a total variational filter and a fourth-order filter. *J. Vis. Commun. Image. R.*, **18**:4 (2007), 322–330.

[29] H. Li, BS Manjunath, and S.K. Mitra. Multisensor image fusion using the wavelet transform. *Graph. Model. Im. Proc.*, **57**:3 (1995), 235–245.

[30] J. Liang, Y. He, D. Liu, and X. Zeng. Image fusion using higher order singular value decomposition. *IEEE Trans. Image Process.*, **21**:5 (2012), 2898–2909.

[31] M. Lysaker and X.C. Tai. Iterative image restoration combining total variation minimization and a second-order functional. *Int. J. Comput. Vision*, **66**:1 (2006), 5–18.

[32] A.G. Mahyari and M. Yazdi. Panchromatic and multispectral image fusion based on maximization of both spectral and spatial similarities. *IEEE Trans. Geosci. Remote*, **49**:6 (2011), 1976–1985.

[33] M.K. Ng and W. Wang. A total variation model for retinex. *SIAM. J. Imaging. Sci.*, **4**:1 (2011), 345-365.

[34] G. Pajares and J. Manuel de la Cruz. A wavelet-based image fusion tutorial. *Pattern Recogn.*, **37**:9 (2004), 1855–1872.

[35] V.S. Petrovic and C.S. Xydeas. Gradient-based multiresolution image fusion. *IEEE Trans. Image Process.*, **13**:2 (2004), 228–237.

[36] G. Piella. Image fusion for enhanced visualization: a variational approach. *Int. J. Comput. Vision*, **83**:1 (2009), 1–11.

[37] G. Piella and H. Heijmans. A new quality metric for image fusion. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 3, pages III–173. IEEE, 2003.

[38] C. Pohl and JL Van Genderen. Multisensor image fusion in remote sensing: concepts, methods and applications. *Int. J. Remote Sens.*, **19**:5 (1998), 823–854.

[39] G. Qu, D. Zhang, and P. Yan. Information measure for performance of image fusion. *Electron. Lett.* , **38**:7(2002), 313–315.

[40] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, **60**:1-4 (1992), 259–268.

[41] P. Scheunders and S. De Backer. Fusion and merging of multispectral images with use of multiscale fundamental forms. *J. Opt. Soc. Amer. A, Optics, Image Science, and Vision*, **18**:10 (2001), 2468–2477.

[42] S. Setzer. Operator splittings, bregman methods and frame shrinkage in image processing. *Int. J. Comput. Vision*, **92**:3 (2011), 265–280.

[43] D.A. Socolinsky and L.B. Wolff. Multispectral image visualization through first-order fusion. *IEEE Trans. Image Process.*, **11**:8 (2002), 923–931.

[44] T. Stathaki. *Image fusion: algorithms and applications.* Academic Press, 2008.

[45] W.W. Wang, P.L. Shui, and X.C. Feng. Variational models for fusion and denoising of multifocus images. *IEEE Signal. Proc. Let.*, **15** (2008), 65–68.

[46] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, **13**:4 (2004), 600–612.

[47] J. Weickert. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998.

[48] C. S. Xydeas and V. Petrovic. Objective image fusion performance measure. *Electron. Lett.*, **36**:4 (2000), 308–309.

[49] Q. Zhang and B. Guo. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Process.* , **89**:7 (2009), 1334–1346.