International Journal of Remote Sensing
Vol. 34, No. 7, 10 April 2013, 2470–2486

Taylor & Francis
Taylor & Francis Group

A variational method for multisource remote-sensing image fusion

Faming Fang[a], Fang Li[b], Guixu Zhang[a]*, and Chaomin Shen[a]

[a]*Department of Computer Science, East China Normal University, Shanghai, China;* [b]*Department of Mathematics, East China Normal University, Shanghai, China*

(*Received 18 February 2012; accepted 8 August 2012*)

With the increasing availability of multisource image data from Earth observation satellites, image fusion, a technique that produces a single image which preserves major salient features from a set of different inputs, has become an important tool in the field of remote sensing since usually the complete information cannot be obtained by a single sensor. In this article, we develop a new pixel-based variational model for image fusion using gradient features. The basic assumption is that the fused image should have a gradient that is close to the most salient gradient in the multisource inputs. Meanwhile, we integrate the inputs with the average quadratic local dispersion measure for the purpose of uniform and natural perception. Furthermore, we introduce a split Bregman algorithm to implement the proposed functional more effectively. To verify the effect of the proposed method, we visually and quantitatively compare it with the conventional image fusion schemes, such as the Laplacian pyramid, morphological pyramid, and geometry-based enhancement fusion methods. The results demonstrate the effectiveness and stability of the proposed method in terms of the related fusion evaluation benchmarks. In particular, the computation efficiency of the proposed method compared with other variational methods also shows that our method is remarkable.

1. Introduction

During the past two decades, with the increasing availability of multisource image data from Earth observation satellites, various imaging sensors have been developed (see Pohl and Genderen 1998; Stathaki 2008). Many image processing tasks, such as image segmentation, feature extraction, and pattern recognition, need as much information of multisource data as possible, while a single sensor may not meet the requirement. Therefore, it is meaningful to integrate different data, which describe the same scene but contain different features, into a new single image that contains all relevant information of the original image set. Multisource image fusion, as a vital technology, can deal with this merging task.

Multisource image fusion is a subarea of a more general topic of data fusion (see Cvejic, Seppanen, and Godsill 2009). It has been a mature and promising field of research in image processing and remote sensing. It provides an effective method to enhance performance for many tasks by combining information from a series of imaging sensors with different modalities. There are many advantages for multisource image fusion: extended range, increased confidence, decreased uncertainty, improved reliability, and robustness of the system performance (see Cvejic, Seppanen, and Godsill 2009).

*Corresponding author. Email: gxzhang@cs.ecnu.edu.cn

ISSN 0143-1161 print/ISSN 1366-5901 online

http://dx.doi.org/10.1080/01431161.2012.744882
http://www.tandfonline.com

In the last decades, many methods and software tools for multisource image fusion have been proposed. According to the different levels of fusion, the existing fusion methods can be roughly divided into three categories: pixel-level fusion, feature-level fusion, and decision-level fusion (see Varshney 1997). In pixel-level fusion, original sensor observations from each sensor are combined directly. In feature-level fusion, features are extracted from observations and are then fused. In decision-level fusion, each sensor makes decisions on its own data first, and these decisions are then combined to produce the final result. Since the results of feature-level and decision-level fusion are less accurate due to the loss of information (see Varshney 1997), we will focus on pixel-level fusion.

Since the aim of image fusion is to integrate the fundamental information of original data into a single complete image, the basic assumption is that this fundamental information is the prominent features and is related to some easily expressed measures, such as image gradient, high-frequency components of a Laplacian pyramid (LP), and the coefficients of wavelet transform (see Zheng et al. 2007). In order to extract prominent features efficiently, many effective methods have been developed during the past few decades. These existing methods include the LP (Burt and Adelson 1983), filter-subtract-decimate pyramid (FSDP) (Anderson 1987), gradient pyramid (GP) (Burt and Kolczynski 1993), discrete wavelet transform (DWT) with Daubechies Spline (DBSS(2,2)) wavelet (DWTD) (Mallat 1989, Ranchin and Wald 1993), shift invariant DWT (SIDWT) with Haar wavelet (Rockinger 1997), and morphological pyramid (MP) (Laporterie and Flouzat 2003).

Recently, the variational method has been widely studied in image processing. The aim of the variational method is to build an energy functional for a certain problem whose minimum is related to the desired result. Compared with other schemes, the variational method has remarkable advantages in both theory and implementation (see Chan, Shen, and Vese 2003). For more information on the variational method, one may refer to Chan, Shen, and Vese (2003), Sapiro (2001), Aubert and Kornprobst (2009), and references therein.

As a variational method, the geometry-based enhancement fusion (GBEF) model (Piella 2009) uses the structure tensor to describe the geometry of all original images. First, it obtains the target structure tensor with geometry-based merging. Then, in order to obtain a more natural and sharp appearance, it performs a variational approach that is a combination of geometry merging, and intensity correction as well as perceptual local contrast enhancement.

In this article, we propose a variational method based on the gradient features for multisource image fusion. The gradient feature method used here was developed by Piella (2009) and Dizenzo (1986). The main idea is that the fused image should have a gradient that approximates the most salient gradient obtained from the multisource original images. At the same time, to make the fused image perceptually more uniform and natural, we thus propose to combine the inputs with the average quadratic local dispersion measure. In particular, the dispersion approach is based on the perceptual colour correction technique proposed by Bertalmio et al. (2007). Furthermore, to tackle the l_1-norm more efficiently, we perform a split Bregman algorithm to solve the proposed model. Note that our method is inspired by the GBEF technique. Compared with GBEF, our contribution lies on (1) a complicated term, i.e. the perceptual contrast enhancement term, is removed in our algorithm, (2) the l_1-norm is used in our method, and (3) a fast algorithm using the split Bregman is adopted to implement our algorithm.

2. The proposed fusion method

The aim of the proposed method is to obtain a high-quality image by integrating the information of the original inputs. Compared with the source images, the fused result should be more informative which is more suitable for vision and applications. Particularly, we can use the visually salient features, such as intensity change, texture, and contour, to measure the information of an image. Since the intensity change is much easier to capture from an image, many methods turn to use it as the feature and merge the feature into the fused image.

2.1. Merging of gradient features

We first introduce some notations that will be used later. Let $u : \Omega \to [0, 1]$ be a greyscale image, where $\Omega \in \mathbb{R}$ represents the rectangle image domain. For a pixel $x = (x_1, x_2) \in \Omega$, $u(x)$ is the intensity at x.

For a greyscale image u, the intensity change is usually expressed by the measure of the gradient $\nabla u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right)^T$. Therefore, $|\nabla u|$ gives the change size, while $\frac{\nabla u}{|\nabla u|}$ is the direction of the change, where $|\cdot|$ is the length of a vector.

For the purpose of fusion, we should obtain the gradients of source images by an appropriate method and combine the gradients into a target gradient $\boldsymbol{g}$. Then, we can obtain the fused image whose gradient ∇u is very similar to $\boldsymbol{g}$. The essential question is how to integrate the gradients. Piella (2009) uses a geometry-based merging method to achieve this aim. However, this method is relatively complicated with low computation efficiency. A straightforward and effective idea is to composite the gradients of all original images $u_n : \Omega \to [0, 1]$, $n = 1, \ldots, N$ with the linear method, i.e. $\boldsymbol{g} = \sum_{n=1}^{N} w_n \nabla u_n$, where w_n is a weight of ∇u_n.

A variety of methods are performed to obtain w_n. We use the following equation to produce a more accurate fused image:

$$w_n = \frac{|\nabla u_n|}{\sum_{i=1}^{N} |\nabla u_i|}. \tag{1}$$

After obtaining the target gradient $\boldsymbol{g}$, we then construct a single image, i.e. the fused image, whose gradient is similar to $\boldsymbol{g}$. Many methods (e.g. Piella 2009) attempt to minimize the quadratic dispersion measure,

$$Q_{l_2} = \int_{\Omega} |\nabla u(x) - \boldsymbol{g}(x)|^2 \mathrm{d}x. \tag{2}$$

An alternative option is to use the l_1 norm. As is known, the l_1 norm is a suitable metric that leads to an l_1-regularized problem, which is common to imaging science. The remarkable advantages of the l_1 norm are that it can achieve an effective result and in turn allows some fast algorithms, such as the primal-dual algorithm (see Chambolle and Pock 2011), augmented Lagrangian method (see Tai and Wu 2009), and split Bregman iteration (see Goldstein and Osher 2009). We try to use the following l_1 formula in proposed method:

$$Q_{l_1} = \int_{\Omega} |\nabla u(x) - \boldsymbol{g}(x)| \mathrm{d}x. \tag{3}$$

The fused image of model (3) may exceed the range [0, 1], so we should rescale the result by adding a range constraint $u(x) \in [0, 1]$.

2.2. Proposed variational functional

For N original input images $u_n : \Omega \to [0, 1]$, $n = 1, \ldots, N$, we aim to produce a composite image $u : \Omega \to [0, 1]$ which is perceptually uniform and natural while keeping the local salient information from all inputs. For this purpose, we combine the average quadratic local dispersion measure (see Bertalmio et al. 2007) with our dispersion measure in (3). The new functional is shown as follows:

$$E = Q_{l_1} + D, \tag{4}$$

where $D = \int\int_\Omega \omega(x, y)(u(y) - \overline{u}^{\omega}(x))^2 dxdy$, $\omega(x, y)$ is a symmetric weight for all $x, y \in \Omega$ with $\int_\Omega \omega(x, y)dy = 1$ for all $x \in \Omega$. $\overline{u}^{\omega}(x)$ is a local average that has two choices (see Bertalmio et al. 2007).

- To make the fused image perceptually more uniform, we use the 'grey world' (GW) assumption (see Buchsbaum 1980), i.e.

$$\overline{u}^{\omega}(x) = \frac{1}{2}. \tag{5}$$

  This leads to

$$D_1 = \int_\Omega \left(u(x) - \frac{1}{2}\right)^2 \mathrm{d}x. \tag{6}$$

- For a given image u_0 that we desire to enhance, another choice of $\overline{u}^{\omega}(x)$ is

$$\overline{u}^{\omega}(x) = \int_\Omega \omega(x, y)u_0(y)\mathrm{d}y. \tag{7}$$

  In this case, D can be seen as a control of the local variance of an image. When considering a limit case with $\omega(x, y) = \delta(x - y)$, we can deduce that

$$D_2 = \int_\Omega (u(x) - u_0(x))^2 \mathrm{d}x. \tag{8}$$

In this article, as suggested by Bertalmio et al. (2007), we shall use the linear combination of the above two choices, i.e.

$$D = \frac{\eta}{2}D_1 + \frac{\mu}{2}D_2, \tag{9}$$

where $\eta \geq 0$ and $\mu > 0$ are parameters.

Therefore, by combining (3) and (9), we can obtain the following proposed energy functional,

$$\min_u E(u) = \int_\Omega |\nabla u(x) - \boldsymbol{g}(x)| + \frac{\eta}{2}\left|u(x) - \frac{1}{2}\right|^2 + \frac{\mu}{2}|u(x) - u_0(x)|^2 \mathrm{d}x, \tag{10}$$

subject to $u(x) \in [0, 1]$. Here, u_0 is the initial image. As a suggestion, u_0 can be set as the weighted sum of u_n, $n = 1, \ldots, N$, i.e. $u_0 = \sum_{n=1}^{N} w_n u_n$, where w_n is defined in (1).

We note that, comparing with the GBEF method proposed by Piella (2009), the proposed energy (10) does not contain the local contrast enhancement portion, which is computation complicated, and (10) contains the l_1-norm which allows many effective fast algorithms.

The formula (10) can be solved by iteratively performing gradient descent steps. However, the non-differentiability of the Q_{l_1} brings computational difficulties (see Tai and Wu 2009). To overcome this drawback, different techniques have been developed, such as dual methods (see Chan, Golob, and Mulet 1999, Carter 2001), Chambolle's algorithm (see Chambolle 2004), the augmented Lagrangian method (see Tai and Wu 2009), and split Bregman iteration (see Goldstein and Osher 2009). Since most of these methods are under the same framework, we use the split Bregman algorithm to implement our model.

2.3. Numerical algorithm for the proposed method based on split Bregman theory

We will now apply the split Bregman framework (see Goldstein and Osher 2009) to solve the l_1-regularized optimization problem (10). As the precondition, we can see that the functional (10) is convex.

The key of our method is that it will separate the l_1 and l_2 components of the energy in (10). Rather than considering (10), we will consider the following constrained problem:

$$\min_{u,\boldsymbol{d}} \int |\boldsymbol{d}| + \frac{\eta}{2}\left|u - \frac{1}{2}\right|^2 + \frac{\mu}{2}|u - u_0|^2 \mathrm{d}x, \quad \text{subject to } \boldsymbol{d} = \nabla u - \boldsymbol{g}. \tag{11}$$

Obviously, this problem is equivalent to Equation (10). To solve this problem, we first convert it to an unconstrained problem using the efficient Bregman iteration approach (see Goldstein and Osher 2009).

$$\begin{aligned}(u^{k+1}, \boldsymbol{d}^{k+1}) = \arg\min_{u,\boldsymbol{d}} \int_\Omega |\boldsymbol{d}| + \frac{\eta}{2}\left|u - \frac{1}{2}\right|^2 \\ + \frac{\mu}{2}|u - u_0|^2 + \frac{\lambda}{2}|\boldsymbol{d} - \nabla u + \boldsymbol{g} - \boldsymbol{b}^k|^2 \mathrm{d}x,\end{aligned} \tag{12}$$

$$\boldsymbol{b}^{k+1} = \boldsymbol{b}^k - \boldsymbol{d}^{k+1} + \nabla u^{k+1} - \boldsymbol{g}, \tag{13}$$

where $\boldsymbol{b}$ is an auxiliary vector and $k \in \mathbb{N} \cup \{0\}$. In order to solve (12), we shall perform this minimization efficiently by alternate iteration with u and $\boldsymbol{d}$ separately. The two steps are performed as

$$u^{k+1} = \arg\min_{u} \int_\Omega \frac{\mu}{2}|u - u_0|^2 + \frac{\eta}{2}|u - \frac{1}{2}|^2 + \frac{\lambda}{2}|\boldsymbol{d} - \nabla u + \boldsymbol{g} - \boldsymbol{b}^k|^2 \mathrm{d}x, \tag{14}$$

$$\boldsymbol{d}^{k+1} = \arg\min_{\boldsymbol{d}} \int_\Omega |\boldsymbol{d}| + \frac{\lambda}{2}|\boldsymbol{d} - \nabla u + \boldsymbol{g} - \boldsymbol{b}^k|^2 \mathrm{d}x. \tag{15}$$

The computational time of the split Bregman algorithm mainly depends on the efficiency of (14) and (15), so we should solve them using some effective methods.

Because the subproblem (14) is differentiable, optimality conditions for u^{k+1} are easily obtained. By differentiating with respect to u, we can derive that

$$\mu(u^{k+1} - u_0) + \eta\left(u^{k+1} - \frac{1}{2}\right) - \lambda\mathrm{div}(\boldsymbol{d}^k - \nabla u^{k+1} + \boldsymbol{g} - \boldsymbol{b}^k) = 0. \tag{16}$$

We further write this optimality condition as

$$(\mu + \eta - \lambda\Delta)u^{k+1} = \mu u_0 + \frac{\eta}{2} + \lambda\mathrm{div}(\boldsymbol{d}^k + \boldsymbol{g} - \boldsymbol{b}^k). \tag{17}$$

The fast Fourier transform (FFT) can be used to solve the above problem

$$\mathcal{F}(K)\mathcal{F}(u^{k+1}) = \mathcal{F}(\mathrm{rhs}^k), \tag{18}$$

where

$$K = \mu + \eta - \lambda\Delta, \tag{19}$$

$$\mathrm{rhs}^k = \mu u_0 + \frac{\eta}{2} + \lambda\mathrm{div}(\boldsymbol{d}^k + \boldsymbol{g} - \boldsymbol{b}^k), \tag{20}$$

$\mathcal{F}$ denotes the FFT, $\mathcal{F}^{-1}$ denotes the inverse FFT, Δ denotes the Laplacian operator, and rhs is an auxiliary vector.

Therefore, the closed form solution of u^{k+1} in (14) is deduced as

$$u^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathrm{rhs}^k)}{\mathcal{F}(K)}\right). \tag{21}$$

The solution of subproblem (15) can be given directly by the following soft-thresholding formula:

$$\boldsymbol{d}^{k+1} = \mathrm{shrink}\left(\nabla u^{k+1} - \boldsymbol{g} + \boldsymbol{b}^k, \frac{1}{\lambda}\right), \tag{22}$$

where

$$\mathrm{shrink}(x, \varsigma) = \frac{x}{|x|} \cdot \max(|x| - \varsigma, 0). \tag{23}$$

So far, the solution of proposed method is obtained. Since the constraint $u^{k+1}(x) \in [0, 1]$ holds, we can use the following formula to truncate the value of u^{k+1} which exceeds [0, 1] in each iteration:

$$u^{k+1} = \max(\min(u^{k+1}, 1), 0). \tag{24}$$

Using the above solver, the overall procedure of the proposed method can be shown in Algorithm 1.

Algorithm 1. The overall procedure for the proposed model

Input: The original images $u_n : n = 1, \ldots, N$
Initialize: $u_0 = \sum_{n=1}^{N} w_n u_n$, $u^0 = u_0$ and $\boldsymbol{d}^0 = \boldsymbol{b}^0 = 0$. Fixed μ, η, λ, and tol
while $||u^k - u^{k-1}|| > tol$
$u^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(rhs^k)}{\mathcal{F}(K)}\right)$
$\boldsymbol{d}^{k+1} = \text{shrink}(\nabla u^{k+1} - \sum_{i=1}^{N} w_n \nabla u_i + \boldsymbol{b}^k, \frac{1}{\lambda})$
$\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\nabla u^{k+1} - \sum_{i=1}^{N} w_n \nabla u_i - \boldsymbol{d}^{k+1})$
$u^{k+1} = \max(\min(u^{k+1}, 1), 0)$
end
Output: the fused image u.

Note that we use the forward differences to approximate gradient operator and the backward difference to approximate divergence operator. For convenience, we extend the boundary of images symmetrically.

Figure 1 illustrates the procedure of proposed image fusion on two multisource remotely sensed images provided by Xu, Chen, and Varshney (2011). First, the gradient

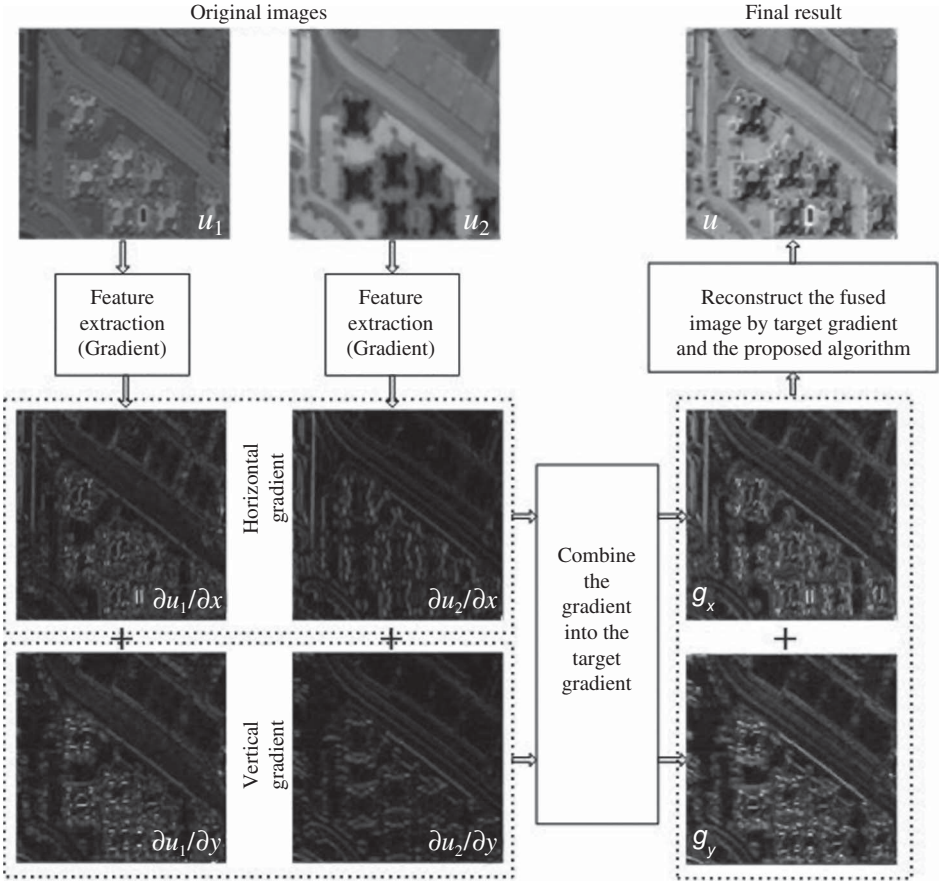

Figure 1. Procedure of multisource image fusion based on the proposed method.

fields $\nabla u_1 = (\partial u_1/\partial x, \partial u_1/\partial_y)^T$ and $\nabla u_2 = (\partial u_2/\partial x, \partial u_2/\partial_y)^T$ are, respectively, produced from the original images u_1 and u_2. Observing the gradient sequence, we can see that the proposed method can extract the salient features from the source image effectively. Second, the gradient sequences are integrated into the target gradient $\boldsymbol{g} = \left[\boldsymbol{g}_x, \boldsymbol{g}_y\right]^T$. The result shows that the integration makes the gradient more informative. As an example, the boundaries of the road are more salient in $\boldsymbol{g}$. Finally, the target gradient is transformed into the fused image u with the proposed functional (10). Observing the source images u_1 and u_2, we can find that the fused image performed by the proposed method is basically an integration of clear contours in the original images. For example, both road and open land are clear in the fused image.

3. Experimental results

In order to show the effectiveness, we test the proposed fusion method on some multisource remotely sensed images. As introduced in Petrovic (2004a), Petrovic provided a multisensor image set that contains 120 pair images (collected by Manchester University and other organizations). This image set includes three types of images: (1) urban, industrial, and natural scenes collected under the USA Airborne Multisensor Pod System (AMPS) programme (Petrovic 2004a), (2) hyper-spectral images of natural scenarios acquired by Bristol University for the UK Defence Research Agency (Brelstaff et al. 1995), and (3) multifocus and extreme exposure image groups (Res 2004). A detailed description of the image set can be found in Petrovic (2004b, 2007) and Zheng et al. (2007). Because our focus is on remotely sensed images, we only use that part in this data set.

All the experiments are implemented in Matlab 7.12 and run on an Intel® 2.33 GHz machine with 2 GB RAM, and all images are at a size of 256 × 256. We set $\mu = 0.5$, $\eta = 0.1$, $\lambda = 0.5$, and $tol = 10^{-6}$ in our experiments by trial and error. Besides, we have found that the parameters are insensitive for our algorithm.

3.1. Evaluation of the fusion method

Evaluation of the image fusion method is a foundational and challenging task, especially when it is hard to distinguish which of the fused results is better (see John et al. 2007). The existing evaluation techniques can be roughly divided into two categories: qualitative visual analysis and quantitative analysis (see Zheng et al. 2007; Stathaki 2008). While qualitative visual analysis is difficult to implement due to the instability of the human visual system (HVS), quantitative analysis can overcome this drawback by using recognized metrics. Since no universal standard exists for image fusion evaluation, both the qualitative visual and quantitative analyses are considered in this article.

To quantitatively evaluate the fused result, the following five typical evaluation metrics are used.

(1) ***Objective quality fusion measure.*** Given u_n, $n = 1, \ldots, N$ multisource input images, and u output fused image, the objective quality fusion measure Q_W (see Piella and Heijmans 2003; Wang, Sheikh, and Simoncelli 2004; Petrovic 2007; Piella 2009) is defined as

$$Q_W = \sum_{n=1}^{N} \sum_{\mathrm{w} \in \boldsymbol{W}} l_n(\mathrm{w}) Q_0(u_n, u|\mathrm{w}), \tag{25}$$

where $l_n(\mathrm{w}) = \lambda \cdot s(u_n|\mathrm{w}) \cdot \max(s(u_n|\mathrm{w}))$ is a local weight indicating the relative importance of input image u_n in window w, $\boldsymbol{W}$ is the family of all windows, λ is the normalization parameter, and $s(u_n|\mathrm{w})$ is local relevance (for example entropy and contrast) of image u_n within the window w. $Q_0(u_n, u|\mathrm{w}) = \frac{2\overline{u_n}\bar{u}+C_1}{\overline{u_n}^2+\bar{u}^2+C_1} \cdot \frac{2\sigma_{u_n}\sigma_u+C_2}{\sigma_{u_n}^2+\sigma_u^2+C_2} \cdot \frac{\sigma_{u_n,u}+C_3}{\sigma_{u_n}\sigma_u+C_3}$ is the local similarity of fused image u and original image u_n given by the structural similarity index, where $\bar{u}$ is the mean of u, and σ_u^2 and $\sigma_{u_n,u}$ are the variance of u and covariance of u_n,u, respectively. The larger the value of Q_W, the better the fused result.

(2) ***Objective image fusion performance measure.*** The normalized weighted objective image fusion performance measure was proposed by Xydeas and Petrovic (2000). It reflects the quality and precision of perceptually important information obtained from the fusion of input images. The original measure is a criterion of two image fusion performance. We now extend it to a multiple image fusion task, which is defined as

$$Q^F = \frac{\sum_{n=1}^{N} \int_\Omega Q^{u_n,u}(x)\omega^{u_n}(x)\mathrm{d}x}{\sum_{n=1}^{N} \int_\Omega \omega^{u_n(x)}\mathrm{d}x}, \tag{26}$$

where $\omega^{u_n}(x)$ is a weight and $Q^{u_n,u}(x) \in [0, 1]$ is an edge information preservation value. $Q^{u_n,u}(x) = 0$ denotes the complete losing of edge information, and $Q^{u_n,u}(x) = 1$ indicates fusion from u_n to u without any waste of information. For $0 \le Q^F \le 1$, the closer Q^F to 1, the more accurate the fusion is.

(3) ***Mutual information.*** Mutual information (see Qu, Zhang, and Yan 2001) is a measurement of the fused image and each of the inputs, which is defined as

$$\mathrm{MI} = \frac{\sum_{n=1}^{N} M(u_n, u)}{\sum_{i=1}^{N} H(u_i)}, \tag{27}$$

where $M(u_n, u)$ denotes the mutual information between u_n and u, and H is the entropy. Generally, the greater the mutual information, the clearer the image, that is to say, the better the fused image.

(4) ***Average gradient.*** Average gradient is sensitive to subtle details of the image. It can be used to evaluate the degree of ambiguity in the image and is calculated as (see Liu et al. 2006).

$$\mathrm{AG} = \frac{1}{|\Omega|} \int_\Omega \sqrt{\left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2}\,\mathrm{d}x. \tag{28}$$

Generally, the greater the average gradient, the better the fusion result.

(5) ***Entropy.*** Entropy is an important index to measure the information of images. The definition of entropy can be shown as (see Liu et al. 2006)

$$H(x) = -\sum_{i=0}^{L-1} p_i \log_2 p_i, \tag{29}$$

where p_i is the probability of the ith grey in the image and L is the grey level. The larger the image entropy is, the better the fusion result.

3.2. Visual analysis

A large number of fusion methods have been performed, and many of them achieve promising results. We now visually analyse the fused images obtained by using the proposed, GBEF (Piella 2009), LP (Burt and Adelson 1983), FSDP (Anderson 1987), GP (Burt and Kolczynski 1993), DWTD (Mallat 1989; Ranchin and Wald 1993), SIDWT (Rockinger 1997), and MP (Laporterie and Flouzat 2003) methods. It should be noted that we set all the parameters in their methods according to the authors' suggestions.

Figures 2–4 illustrate the comparison of fused results on multisource remotely sensed images. In each figure, (*a*) and (*b*) represent the source images obtained by different sensors. (*c*)–(*j*) show the resulting images produced by the proposed, GBEF, LP, FSDP, GP, DWTD, SIDWT, and MP methods, respectively. Observing the figures carefully, we can find that the proposed method contains more details than all other methods. For example, in Figure 4(*d*)–(*j*), the building appears blurring and dimming, while in (*c*) it appears visually clear and uniform.

Based on the above observation, we can conclude that, for visual performance, the proposed image fusion method performs slightly better or equivalently to the other methods on multisource remotely sensed images.

3.3. Quantitative analysis

Tables 1–3, which are, respectively, related to Figures 2–4, show the values of objective quality fusion measure Q_W, objective image fusion performance measure Q^F, mutual information MI, average gradient AG, and entropy H for the image fusion schemes including the proposed, GBEF, LP, FSDP, GP, DWTD, SIDWT, and MP methods.

From these tables, the proposed image fusion method achieves the best performance, resulting quantitative analysis indices include Q_W, Q^F, MI, AG, and H. We take Table 1 as an example. The proposed method outperforms all of the other methods in Q_W, Q^F, MI, and H measures. Especially in the Q^F measure, the value of the proposed method is 20 percentage points larger than that of the second best method (GBEF method). Although the AG measure produced by the proposed method is not the best, it is only sightly smaller than MP and greater than all other methods. Similar results can be found in Tables 2 and 3.

As can be seen from the above analysis, the proposed image fusion method is generally better than other conventional fusion schemes such as GBEF, LP, FSDP, GP, DWTD, SIDWT, and MP methods, with respect to the related quantitative fusion analysis indices including Q_W, Q^F, MI, AG, and H, on multisource remotely sensed images.

We also perform the quantitative evaluation for the method which uses the proposed scheme but the Q_{l_1} portion in (10) is replaced with Q_{l_2} in (2). We call this method the proposed fusion method with L2-norm (PFL2). For example, in Table 1, all quantitative analysis indices of the proposed method are larger than that of PFL2. That is, the overall results of the proposed method are better than PFL2. Thus, the results of the proposed method contain more subtle details than PFL2. Therefore, we may draw the conclusion that it is more reasonable for the proposed method to use the l_1-norm over the l_2-norm.

3.4. Computation efficiency analysis

To evaluate the computation efficiency of the proposed method, we further compare it with the GBEF method, which is also a variational approach, using the images from Figures 2–4

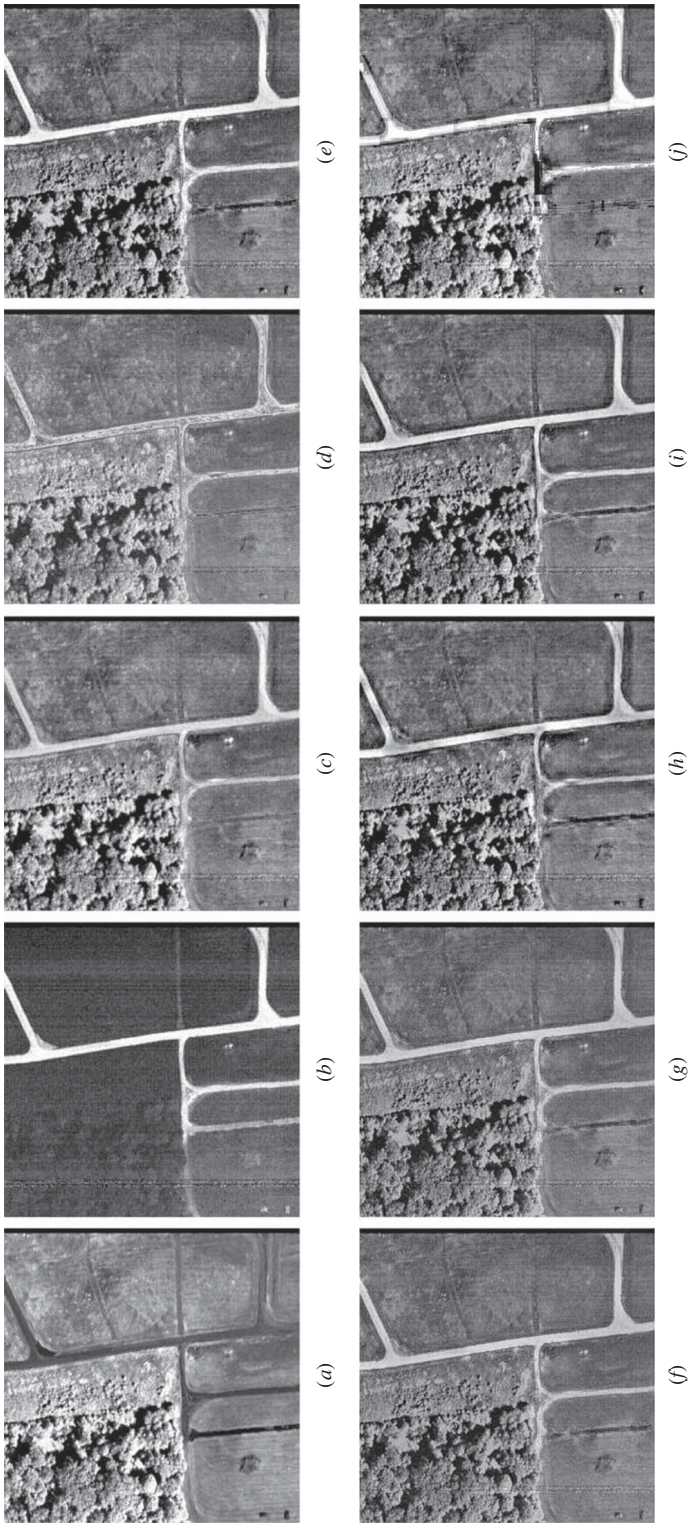

Figure 2. (*a*,*b*) The original multisource remotely sensed images (size: 256 × 256). (*c*) The result of the proposed method. (*d*) The result of the GBEF method. (*e*) The result of the LP method. (*f*) The result of the FSDP method. (*g*) The result of the GP method. (*h*) The result of the DWTD method. (*i*) The result of the SIDWT method. (*j*) The result of the MP method.

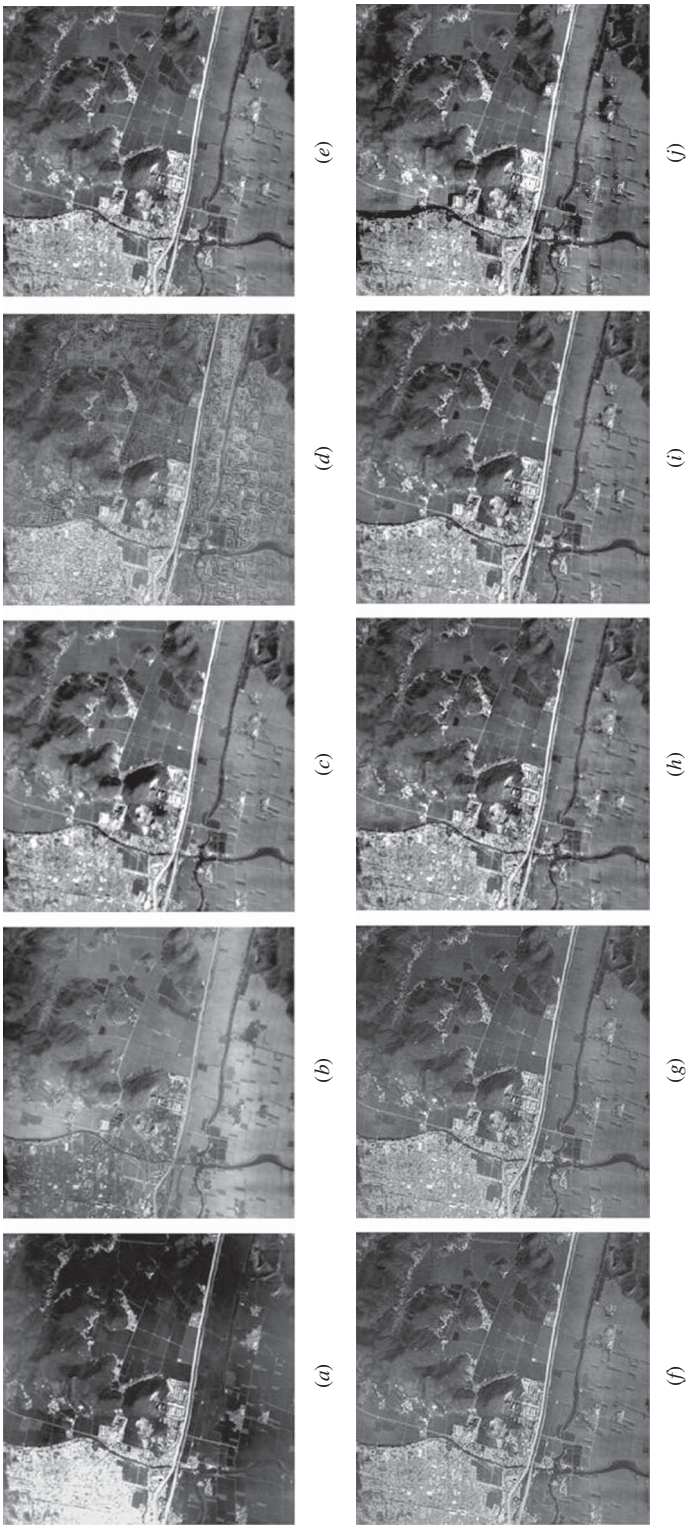

Figure 3. (*a*,*b*) The original multisource remotely sensed images (size: 256 × 256). (*c*) The result of the proposed method. (*d*) The result of the GBEF method. (*e*) The result of the LP method. (*f*) The result of the FSDP method. (*g*) The result of the GP method. (*h*) The result of the DWTD method. (*i*) The result of the SIDWT method. (*j*) The result of the MP method.

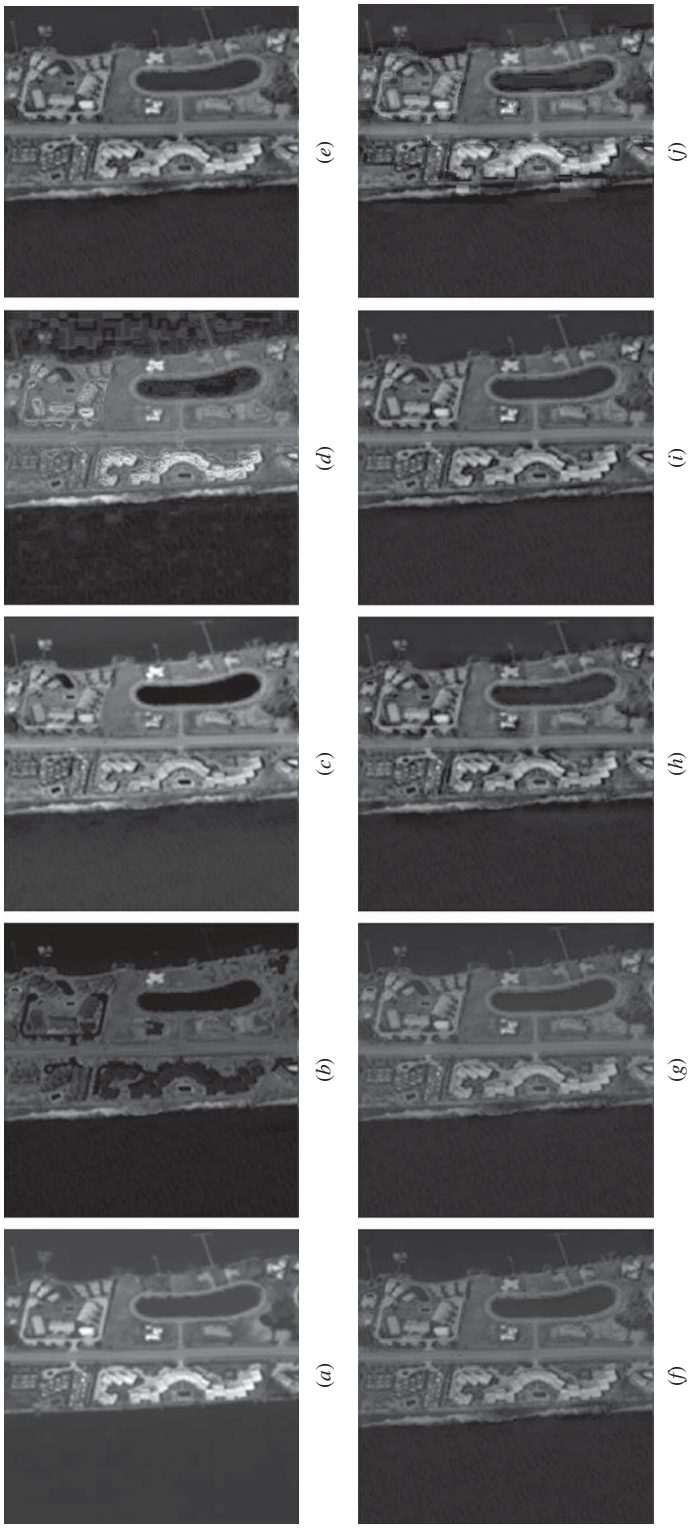

Figure 4. (*a*,*b*) The original multisource remotely sensed images (size: 256 × 256). (*c*) The result of the proposed method. (*d*) The result of the GBEF method. (*e*) The result of the LP method. (*f*) The result of the FSDP method. (*g*) The result of the GP method. (*h*) The result of the DWTD method. (*i*) The result of the SIDWT method. (*j*) The result of the MP method.

Table 1. Comparison of fused results in Figure 2.

Methods	Q_W	Q^F	MI	AG	H
Proposed	**0.6193**	**0.5246**	**0.1705**	21.6480	**7.4163**
PFL2	0.5931	0.4242	0.1481	14.3985	6.8397
GBEF	0.4830	0.4407	0.1515	18.0999	7.3993
LP	0.4609	0.4140	0.1216	20.5918	7.3765
FSDP	0.4972	0.3891	0.1180	17.5436	7.2338
GP	0.5031	0.3945	0.1188	17.2778	7.2269
DWTD	0.4279	0.3784	0.1043	21.5563	7.3787
SIDWT	0.4606	0.4182	0.1115	20.8357	7.3585
MP	0.3967	0.3537	0.1161	**23.5026**	7.4119

Note: The bold values are the best.

Table 2. Comparison of fused results in Figure 3.

Methods	Q_W	Q^F	MI	AG	H
Proposed	**0.6409**	**0.4647**	**0.1469**	16.0696	**7.5352**
PFL2	0.6143	0.3399	0.1156	10.1768	6.5246
GBEF	0.4407	0.3666	0.1383	15.6350	7.4686
LP	0.5771	0.4542	0.1369	17.4117	7.5254
FSDP	0.6031	0.4336	0.1425	13.7953	7.2813
GP	0.6121	0.4410	0.1468	13.5528	7.2721
DWTD	0.5535	0.3993	0.1266	17.6125	7.4892
SIDWT	0.5928	0.4577	0.1407	16.8125	7.4720
MP	0.4901	0.4078	0.1418	**19.0126**	7.4962

Note: The bold values are the best.

Table 3. Comparison of fused results in Figure 4.

Methods	Q_W	Q^F	MI	AG	H
Proposed	**0.7020**	**0.4977**	0.3104	**7.9930**	**6.6936**
PFL2	0.6761	0.4832	0.3420	7.6907	6.6933
GBEF	0.5994	0.4632	**0.3567**	7.8131	6.6806
LP	0.4824	0.4732	0.1950	7.3325	6.2461
FSDP	0.5308	0.4634	0.1622	5.2999	6.0871
GP	0.5370	0.4734	0.1488	5.2396	6.0721
DWTD	0.4693	0.4390	0.1470	7.4169	6.2083
SIDWT	0.4877	0.4881	0.1775	7.2057	6.1149
MP	0.3901	0.4087	0.1496	7.9492	6.2920

Note: The bold values are the best.

with respect to time costs. Besides, the computational time of the PFL2 method is also presented. The results are shown in Table 4. All of the values are average time obtained by repeatedly processing 20 times for each figure.

We can see from Table 4 that the time costs of the proposed method are almost one-fifth those of the GBEF method, and less than half those of the PFL2 method. The main reason is that a time-consuming component used for local enhancement exists in the GBEF method, and the time variable in the gradient descent flow influences the speed of the PFL2 method (here we use the gradient descent flow to implement the PFL2), while the proposed method has no complex partition and has used the split Bregman algorithm.

Table 4. Comparison of computation time of different methods (in seconds).

Methods	Figure 2	Figure 3	Figure 4
GBEF	5.5765	4.4198	4.5968
PFL2	2.7475	2.0297	1.9886
Proposed	**0.9652**	**0.7977**	**0.8395**

Note: The bold values are the best.

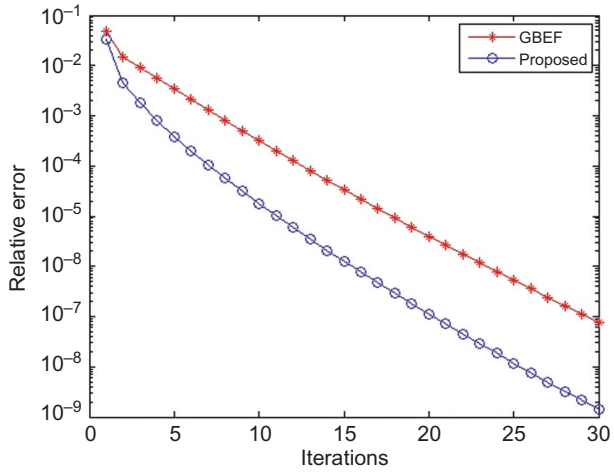

Figure 5. Error versus iterations for the proposed method and GBEF method (Piella 2009). The error at iteration k is defined as $\left\|u^k - u^{k-1}\right\| / \left\|u^{k-1}\right\|$, where u^k is the approximation at iteration k. The convergence result is for the images in Figure 2.

The convergence speed of the proposed method is also presented in Figure 5, where as we have plotted the error versus iterations for the images in Figure 2. To compare our method to another method, we also draw the convergence curve for the GBEF method. Observing the result, we can find that the proposed method converges faster than the GBEF method. In addition, the proposed method has reached promising results after the first 10 iterations, while the GBEF method needs at least 20 iterations to achieve the same effect.

4. Conclusions

We have introduced an image fusion method for multisource remotely sensed images. The proposed method (1) extracted the gradient features from original images, (2) integrated them into a new target gradient, and (3) obtained the fused image by minimizing an energy functional. For tackling the functional more efficiently, we used the split Bregman algorithm to obtain the final solution. Compared with the GBEF technique, our method is simple since it does not contain the complicated perceptual contrast enhancement term. Besides, our method uses the l_1-norm, which allows the fast algorithm (split Bregman iterator). In order to prove the effectiveness of the proposed method, we introduced the qualitative visual and quantitative analyses. The comparing results demonstrated that our method is generally better than many existing fusion schemes with respect to the related evaluation criteria. In particular, the comparison in time efficiency with the GBEF and PFL2 methods

also demonstrated the effectiveness of the proposed method. Therefore, we can draw a conclusion that the proposed method can effectively extract salient features and composite it into the resulting image.

The feature extraction and integration of multisource images is a developing problem in multisource remote-sensing image processing. Our method is effective, but it also can be improved in terms of feature extraction and integration functional. We will further extend our model for more accurate fusions.

Acknowledgements

This work was supported by the 973 Programme (2011CB707104), the National Science Foundation of China (Nos. 11001082, 10971066, and 61273298), and the Programme for New Century Excellent Talents in University (No. NCET-08-0193).

References

Anderson, C. 1987. "A Filter-Subtract-Decimate Hierarchical Pyramid Signal Analyzing and Synthesizing Technique." U.S. Patent 718 104.

Aubert, G., and P. Kornprobst. 2009. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147. Berlin: Springer-Verlag, Applied Mathematical Sciences.

Bertalmio, M., V. Caselles, E. Provenzi, and A. Rizzi. 2007. "Perceptual Color Correction through Variational Techniques." *IEEE Transactions on Image Processing* 16: 1058–72.

Brelstaff, G., C. Prraga, T. Troscianko, and D. Carr. 1995. "Hyperspectral Camera System: Acquisition and Analysis." In *Proceedings of the SPIE Conference on Geographic Information Systems, Photogrammetry and Geological/Geophysical Remote Sensing*, edited by J. B. Lurie, J. J. Pearson, and E. Zilioli, vol. 2587, 150–9. Paris: SPLE.

Buchsbaum, G. 1980. "A Spatial Processor Model for Object Color Perception." *Journal of the Franklin Institute* 310: 1–26.

Burt, P., and E. Adelson. 1983. "The Laplacian Pyramid as a Compact Image Code." *IEEE Transactions on Communications* 31: 532–40.

Burt, P., and R. Kolczynski. 1993. "Enhanced Image Capture Through Fusion." In *Proceedings of the 4th International Conference on Computer Vision*, edited by B. Triggs, A. Zisserman, and R. Szeliski, 173–82. Los Alamitos, CA: IEEE Computer Society Press.

Carter, J. 2001. "Dual methods for Total Variation-Based Image Restoration." PhD thesis, University of California, Los Angeles.

Chambolle, A. 2004. "An Algorithm for Total Variation Minimization and Applications." *Journal of Mathematical Imaging and Vision* 20: 89–97.

Chambolle, A., and T. Pock. 2011. "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging." *Journal of Mathematical Imaging and Vision* 40: 120–45.

Chan, T., G. Golub, and P. Mulet. 1999. "A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration." *SIAM Journal on Scientific Computing* 20: 1964–77.

Chan, T., J. Shen, and L. Vese. 2003. "Variational PDE Models in Image Processing." *Notices of the American Mathematical Society* 50: 14–26.

Cvejic, N., T. Seppanen, and S. Godsill. 2009. "A Nonreference Image Fusion Metric Based on the Regional Importance Measure." *IEEE Journal of Selected Topics in Signal Processing* 3: 212–21.

Dizenzo, S. 1986. "A Note on the Gradient of Multi-Image." *Computer Vision, Graphics, and Image Processing* 33: 116–25.

Goldstein, T., and S. Osher. 2009. "The Split Bregman Method for L1-Regularized Problems." *SIAM Journal on Imaging Sciences* 2: 323–43.

Imagefusion.org. 2004. The Online Resource for Research in Image Fusion. Accessed March 3, 2011. www.imagefusion.org.

John, J., J. Robert, S. Nikolov, D. Bull, and N. Canagarajah. 2007. "Pixel- and Region-Based Image Fusion with Complex Wavelets." *Information Fusion* 8: 119–30.

Laporterie, F., and G. Flouzat. 2003. "The Morphological Pyramid Concept as a Tool for Multi-Resolution Data Fusion in Remote Sensing." *Journal Integrated Computer-Aided Engineering* 10: 63–79.

Liu, W., J. Huang, and Y. Zhao, 2006. "Image Fusion Based on PCA and Undecimated Discrete Wavelet Transform." In *Neural Information Processing*, vol. 4233, 481–8. Berlin: Springer.

Mallat, S. 1989. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 674–93.

Petrovic, V. 2004a. *The Image Collection of Manchester University*. Accessed April 16, 2011. http://www.imagefusion.org/images/petrovic/petrovic.html.

Petrovic, V. 2004b. "Subjective Image Fusion Evaluation Data." *Technical Report*, Imaging Science Biomedical Engineering, University of Manchester.

Petrovic, V. 2007. "Subjective Tests for Image Fusion Evaluation and Objective Metric Validation." *Information Fusion* 8: 208–16.

Piella, G. 2009. "Image Fusion for Enhanced Visualization: A Variational Approach." *International Journal of Computer Vision* 83: 1–11.

Piella, G., and H. Heijmans. 2003. "A New Quality Metric for Image Fusion." *Proceedings of the IEEE International Conference on Image Processing*, 3: 173–6.

Pohl, C., and J. Genderen. 1998. "Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications." *International Journal of Remote Sensing* 19: 823–54.

Qu, G., D. Zhang, and P. Yan. 2001. "Medical Image Fusion by Wavelet Transform Modulus Maxima." *Journal of the Optical Society of America* 9: 184–90.

Ranchin, T., and L. Wald. 1993. "The Wavelet Transform for the Analysis of Remotely Sensed Images." *International Journal of Remote Sensing* 14: 615–19.

Rockinger, O. 1997. "Image Sequence Fusion Using a Shift-Invariant Wavelet Transform." *Proceedings of the IEEE International Conference on Image Processing*, 3: 288–91.

Sapiro, G. 2001. *Geometric Partial Differential Equations and Image Processing*. Cambridge: Cambridge University Press.

Stathaki, T. 2008. *Image Fusion: Algorithms and Applications*. Amsterdam: Elsevier.

Tai, X., and C. Wu. 2009. "Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model." In *Proceedings of the Scale Space and Variational Methods in Computer Vision*, edited by X.-C. Tai, K. Morken, M. Lysaker, and K.-A. Lie, vol. 5567, 502–13. Berlin Heidelberg: Springer.

Varshney, P. 1997. "Multisensor Data Fusion." *Electronics and Communication Engineering Journal* 9: 245–53.

Wang, Z., A. Sheikh, and E. Simoncelli. 2004. "Image Quality Assessment: From Error Measurement to Structural Similarity." *IEEE Transactions on Image Processing* 13: 600–13.

Xu, M., H. Chen, and K. Varshney. 2011. "An Image Fusion Approach Based on Markov Random Fields." *IEEE Transactions on Geoscience and Remote Sensing* 49: 5116–27.

Xydeas, C., and V. Petrovic. 2000. "Objective Image Fusion Performance Measure." *Electronics Letters* 36: 308–9.

Zheng, S., W. Shi, J. Liu, G. Zhu, and J. Tian. 2007. "Multisource Image Fusion Method Using Support Value Transform." *IEEE Transactions on Image Processing* 16: 1831–9.