Numer. Math. Theor. Meth. Appl. Vol. **17**, No. 2, pp. 275-309
doi: 10.4208/nmtma.OA-2023-0091 May 2024

Fusing Infrared and Visible Images via a First-Order Model

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

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

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

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

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

---

1. Introduction

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

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

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

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

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

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

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

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