Pansharpening via Deep Guided Filtering Network

Ahmad Alsmadi*, Shuyuan Yang*, Kai Zhang*

Dep. of Electrical Engineering, Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China
Xi’an, China

ahmadsmadil@yahoo.com, syyang2009@gmail.com, 2428902470@gmail.com

Abstract—In order to enhance spatial details and reduce spectral distortions in the fused image, in this paper, we introduce a simple pan-sharpening way based on deep guided filtering network, our proposed method was designed by combining the deep guided filtering network with intensity hue saturation (IHS) transform. First, by using the intensity component (I) which is calculated from the up-sampled MS image via IHS transform; second, a deep guided filtering network is constructed to filter the panchromatic (PAN) image to extract a semantic detail map from the panchromatic image; third, the details are injected into every band of the up-sampled multi-spectral image to acquire the fused image by an adaptive injection, which will aid in controlling the quantity of the injected details. We empirically evaluated the proposed method under a several data-sets from the GEO-2, IKONOS-2, and GF-2 satellites, and the results are compared with those of its counterparts. The results demonstrate that the proposed method can well-preserve spatial and spectral information, in both subjective and objective aspects. The results show that the proposed approach can acquire fusion results and outperform the other methods in-indices of SAM, RMSE, ERGAS, SAM, CC, UIQI and Q4.

Keywords—Pansharpening, intensity-hue-saturation (IHS) transform, guided filtering network.

I. INTRODUCTION

The target of image fusion is to generate a single image by combining relevant information from two or more images, therefore, satellites-remote-sensing image fusion has been a hot research topic of image fusion [1]. There are many commercial satellites such as GEO-2, IKONOS-2, and GF-2, that provides a high-spatial/low-spectral resolution panchromatic (PAN) image and a high-spectral/low-spatial resolution multi-spectral (MS) image concurrently. To recover this problem, there is an effective way to combine a multi-spectral and a panchromatic images, which is called remote-sensing image fusion, intended for a result in a single image containing both a higher spatial/spectral resolutions [2]. Today the fused remote sensing image is a great significance technique and is used in providing the imagery seen in the common Google-Maps/Earth and Microsoft-Bing-Maps products. There are also some applications within the field of remote-sensing that benefit from pan-sharpened imagery, such as change-detection, classification [3], vegetation identification, and lithology-analysis [4].

In the last two decades, many fusion approaches have been published to improve the spatial-resolution of the LR multispectral image. The main goal of these approaches is to inject the spatial-detail-information, which is extracted from the panchromatic image into the multi-spectral image within a predefined process, which is the injection gains. However, to improve the spatial resolution of LR band, there are two common methods, which are the component-substitution (CS) based method and the multi-resolution-analysis (MRA) based method [5]. The component-substitution-based method is most commonly used in remote-sensing image fusion, such as (IHS) transform [6]-[7], Gram-Schmidt (GS) [8] and principal component-analysis (PCA) [9]-[10]. The fusion images during this method suffer from spectral distortions.

In the multi-resolution-analysis (MRA) based method, such as the Laplacian-pyramid (LP) [11] and the “a trous” wavelet transform (ATWT) [12]. In order, the MRA based method can provide more accurate spectral information but suffer from spatial distortions. In the predefined process, the quantity of panchromatic image spatial-detail-information that is injected into the multi-spectral image is determined by the injection gains, which affects the fusion result immediately [13]. Recently, Remote-Sensing-Image-Fusion Based on Adaptive IHS and Multi-scale Guided Filter (AIHS) has been proposed in [14]. The method in [14] takes advantage of the multi-scale form of filtering, in order to extract sufficient details from the panchromatic image. Hence, the obtained fused results through this method are impressive. They used the panchromatic image as the input image of a guided filter and the (I) component of the multi-spectral image was used as the guided image at every stage. Li et al. [15] proposed a multi-stage guided filter based pan-sharpening method (MGF). They selected high-spatial resolution panchromatic image as the guidance in transferring the structures into the up-scaled MS images and smoothing the PAN image itself. Qi et al. [16] also proposed multi-stage guided filtering based on hyperspherical color transform (HCS), they used the (I) component of the multi-spectral image as the input of the guided filter, which was extracted from the hyperspherical color transformation (HCT), and the PAN image was used as the guidance-image at the first stage, but in the second stage, at the same time, they made the PAN image as the guidance and the filtering input. This method obtains the fused image with a high-spatial-resolution, but the obvious color distortions often occur in the sharpening process. Note that, this method is more effective for WorldView-2 satellite imagery, which utilizes more bands.

In this paper, we propose a novel pan-sharpening approach. By using the guided filter, the approximation version of the
panchromatic image is obtained, in which the input image is the panchromatic image in the first stage of guided filter, and the (I) component of the multi-spectral image is used as the guidance image, because the (I) component change has a bit effect on the spectral-information and due to its ease of usage. In the second stage, the opposite of the aforementioned steps will be executed. Moreover, to control the detail information injected into the multi-spectral image, an adaptive improved injection gains procedure is designed. We employ a guided filter to filter the panchromatic image and the (I) component of the multispectral image to obtain the approximation low-resolution image in the first stage and the second stage, respectively. The guided filter calculates the filtering output by taking into consideration the content of a guidance-image, as it can transfer the structures of the guidance-image to the filtering output [17]-[18]. We use a multi-stage strategy, in order to be able to extract more detail information. Hence, the intensity component is obtained by using linear weighting. Enhancing the efficiency of our method can be achieved by the improving the injection gains that are being obtained from both of the panchromatic and each band of the multi-spectral images. Super resolution(SR) reconstruction was accomplished by employing an iterative-back-projection (IBP), which is become a known and computationally efficient method for improving the spatial-resolution of the image [19]. However, IBP algorithm has some presenting limitations, like ringing artifacts in the strong edge-area of the image [20]. In this method, the IBP is being used for each band in the MS image.

Our fusion method produces a higher spatial and spectral resolution with minimum-spectral-distortion for the fused image among of existing methods. The contributions of our method are synthesis, in which the (I) component is obtained from the upsampled MS image, a multi-stage guided filter procedure to filter the panchromatic image, which extracts more detail information, as a semantic detail map is injected into each band to obtain the fused image by a model-based procedure, then, it makes an iterative back projection (IBP) for each band. For both of the subjective and the objective evaluations, the experimental results indicate that our method has best results among of other existing methods.

The rest of this paper is structured into five sections. In Section II, briefly reviews the guided filter. In Section III our proposed fusion method is described. We apply our proposed method on different data-sets and compare our fusion results with some other existing method in Section IV. In Section V, a conclusion is presented.

II. GUIDED FILTER

As we know, the edge preserving filters have been recently a hot topic in image processing, which has received a great attention from the research community. In order to avoid ringing artifacts, there are some of edge preserving smoothing filters such as guided filter [21], weighted-least-squares [22], and bilateral-filter [23]. Currently, the fastest edge preserving filter is the guided filter. It is useful for detail enhancement, image matting/feathering, and dehazing. The image guided filter function implements edge preserving smoothing on an image, by using the content of a second image, which is being used as a guided image, in order to effect the filtering. Furthermore, the guided image could be the same image. Like other filtering processes, the guided filter is a neighborhood process, but takes consideration while computing the value of the output pixel; the statistics of a region in the corresponding spatial neighborhood in the guidance-image [24]. The guided filter can avoid gradient reversal artifacts. In this paper, the guided filter is employed for pan-sharpening. It is worth to mention; the guided filter is a local linear model between filtering output image O and the guidance image Y. Theoretically, by using a guided image Y in filtering the input image I, in order to get the output image O, moreover, the output image O has the structures of the guidance-image, and it can maintain the main information of the input image. Assuming that O is a linear transform of Y in a window w_k centered at a pixel k, it can be calculated by the following equation:

\[ O_i = a_i Y_i + b_i \forall i \in w_k \]  \hspace{1cm} (1)

where i is the pixel index, and w_k is a squared window of size \((2r +1)\times (2r +1)\). The following equation will be used to represent the guided filtering operation for this paper:

\[ O = G_{r,\zeta}(I, Y) \]  \hspace{1cm} (2)

where G denotes the guided filter function, the parameters r and \(\zeta\) are the inputs of guided filter, which are determined the size and the blur degree of the guided filter, respectively. Note that, I and Y are the input and the guidance images, respectively.

III. PROPOSED METHOD

The proposed pan-sharpening method is shown in Fig. 1. It is based on the component-substitution (CS) method and the multiresolution-analysis (MRA) method, which use the popular edge-preserving guided filter and iterative back projection (IBP) in injecting the extracted spatial detail information from the PAN image into the up-sampled MS image. Our method includes two main steps: Step1) using the PAN image to extract the semantic detail map; Step2) computing the injection gain for each band.

A. Extracting a semantic detail map from the panchromatic Image

In order to acquire the spatial-information, the conventional methods estimate the variation between the panchromatic image and the intensity component of multi-spectral image or any type of a low-pass filtered panchromatic image. However, these methods suffer from the spectral-distortion because the spatial information will blend with low-frequency components [22]. We proposed a simple way to extract the semantic detailed
map by estimating the variation between the panchromatic image and the multi-stage guided filter, which aids in overcoming the aforementioned problem. The approximation image of the input image is the result of the guided filter. Hence, the semantic detailed map of the input image is the variation between the input image and the approximation image. In this paper, we present a multi-stage guided filter, in which the I component is used as the guidance image and the panchromatic image as the input image in the first stage (s=1) of guided filter. In the other stages of our proposed method, the panchromatic image is used as the guidance image and the result of previous stage is used as the input image. The procedure of the first stage can be formulated as follows:

$$ GF(PAN) = Gr, \zeta(PAN,I) $$  

(3)

where $GF(PAN)$ denotes the guided filter output. However, for $s_{th}$ stage ($s > 1$), the procedure of the guided filter can be formulated as

$$ GF^s(PAN) = Gr, \zeta(PAN^{(s-1)},PAN) $$  

(4)

where $PAN^{(s-1)}$ is the approximation layer and $(s - 1)_{th}$ represents the index level of guided filtered output. The spatial detail $PAN^s_B$ for the $s_{th}$ stage can be shown as

$$ PAN^s_B = GF^{(s-1)}(PAN) - GF^s(PAN) $$  

(5)

Note that, when $s=1$ the original PAN image is $GF^{(s-1)}(PAN)$.

### B. Computing the injection gain for each band

The injection gains are defined by the quantity of the panchromatic image spatial detail information, which is being injected into the multi-spectral image. Consequently, the injection gain has been shown to be critical in the spatial detail injection model. Therefore, if the quantity of the spatial detail is insufficient, the fused image will not reach its desired limits and if the quantity of the spatial detail was very high, the fused result will suffer from spectral distortion, which will cause redundant information. However, each band in MS image has different characteristics, cause the variety of the spectral reflectance, therefore, each band has different trade-off parameter $\beta$.

Accordingly, in [14], the trade-off parameters can be generated adaptively, by resolving the next optimization problem:

$$ \min_{\beta_1, \ldots, \beta_n} \| w_p - \sum_{i=1}^n \beta_i WM_i \|^2 \text{ s. t. } \beta_1 \geq 0, \ldots, \beta_n \geq 0 $$  

(6)

where $n$ represents the number of the multi-spectral band and $\beta_i$ represents the $i_{th}$ band trade-off parameter of the MS image’s. Thenceforward, we can acquire the improved edge-detecting-weighting-matrix as

$$ g_i = \frac{M_i}{\sum_{i=1}^n M_i} ((1 - \beta_i)w_p + \beta_i w_M) $$  

(7)

where $w_A$ represents the edge-detecting-weighting-matrix of image A

$$ w_A = \exp\left(-\frac{\lambda}{|\nabla A|^4 + \varepsilon}\right) $$  

(8)

where $\nabla A$ denotes a gradient of the image A and $\lambda$ and $\varepsilon$ are the tuning parameters. Fig. 2 illustrates the following procedure obtaining the injection gain.
Fig. 2 The whole procedure of obtaining the injection gain.

Now to obtain the $F$ according to Fig. 1, the total semantic details of the panchromatic image will inject into the up-sampled multi-spectral image as formulated in Equation (9):

$$F_i = M_i + g_i \sum_{s=1}^{K} PAN_D^g$$  \hspace{1cm} (9)

where $K$ is the number of stages for the guided filter.

Then using IBP for each MS band by considering $F_i$ as LR image and the up-sampled MS image as HR image to improve the spatial-resolution of the fusion result. To improve more spatial detail, we present multiple filter stages, but using an excessive filter stage will suffer from some problems in the fused image such as redundant information and produce artifacts while increasing the computational cost. Fig. 3, shows the fusion results of GEO-2 images with different guided filter stages. The normalized results were obtained with respect to the different filter stage as shown in Fig. 5. According to Fig. 5, the best-fused results can be seen in stages 1 and 2.

Fig. 3 Illustration fusion results of GEO-2 images with different guided filter stages. Figures (a) - (d) denote the filter stages 1 to 4, respectively.

IV. EXPERIMENTAL STUDY AND ANALYSIS

To evaluate the performance of our method, we consider Wald’s protocol, which states that a fused image should be have a close similarity to the image that the corresponding sensor would observe at the highest spatial resolution [25]. We are selected the original multi-spectral images as the true images to be compared with the fusion results, therefore, the experiments are accomplished on the degraded images. The parameter settings, which is used in the proposed method are discussed. To show the behavior of the proposed fusion method three data sets i.e., IKONOS-2, GEO-2, and GF-2 are used experimentally. After the subjective evaluation should be appraised the performance of each fusion method quantitatively. Hence, six typical evaluation metrics, namely, the correlation coefficient (CC), the root mean square-error (RMSE), universal-image-quality indexes (UIQI), the erreur-relative-global-adimensionnelle de synth`ese (ERGAS), a quaternion-based coefficient (Q4) index, and the spectral angle-mapper (SAM), are used.

A. Parameters settings

The inputs of the guided filter are critical factors that must be obtained. Moreover, an excessive number of iterations of IBP may lose some part from the fused image. Note that, in our proposed method, the Num-of-iterations is fixed to 5. The experiments in this section are performed on the GEO-2 images. Fig. 4, shows the normalized results with respect to different input parameters of the guided filter, which are the square window of radius $r$ and the regularization parameter $\zeta$. There are three cases of interest: a) $r=2$ and $\zeta = 0.1^2$; b) $r=4$ and $\zeta = 0.2^2$; c) $r=8$ and $\zeta = 0.4^2$. Note that, the quality indicators are calculated and some indexes are normalized to [0 1]. The better-fused results have larger CC, UIQI, and $Q_4$, and smaller RMSE, SAM, and ERGAS. Case (a) is shown in Fig. 4, where the fused result achieves the optimal effect. Note that, the authors in [15] fixed the parameters $r=2$ and $\zeta = 0.001,0.8$ in the first and the second stage, respectively. Moreover, the authors in [16] fixed the parameters $r=4$ and $\zeta = 0.1 \times nbits$, where nbits is the radiometric resolution of each image.

Fig. 4 Performance of our fusion method with respect to different input guided filter parameters.

In this paper, the rest of parameters are set as follows, the multi-spectral image is matched to the same size of the panchromatic...
image, the parameters $\lambda = 10^{-9}$ and $\epsilon = 10^{-9}$ are set as in [14].

We compared our proposed fusion method with several typical fusion methods: the AIHS method [14]; the (MGF) method [15]; the (HCS) method [16].

**B. Experiments with IKONOS-2 Data**

The IKONOS-2 system simultaneously offers a four-band MS image with 4m resolution and a single-band PAN image with 1m resolution. The results for IKONOS-2 are shown in Fig. 6. The size of the PAN image and the MS image in this experiment are $2048 \times 2048$ and $512 \times 512$, respectively. In this study, by using the protocol in [25], we degrade the original MS and PAN images by a factor of 4, then we utilize a simulated IKONOS-2 LR-MS image with the size of $128 \times 128$ and a corresponding panchromatic image with the size of $512 \times 512$. The original multi-spectral image with 4m resolution is used as the true-image shown in Fig. 6 (a), in order to perform comparative studies between the true-image and the fusion result.

In this section, we analyze another group of experimental results of image fusion with GEO-2 data to further clarify the performance of our fusion approach. The GEO-2 data set provides a four band 2.8m resolution multi-spectral image and a 0.7m resolution PAN image. The results for GEO-2 are presented in Fig. 7. The size of the panchromatic and the multi-spectral images in this experiment are $512 \times 512$ and $512 \times 512$, respectively. In this study, by using the protocol in [25], we degrade the original multi-spectral image by a factor of 4, then we utilize a simulated GEO-2 LR-MS image with the size of $64 \times 64$ and a corresponding PAN image with the size of $512 \times 512$. Fig. 7 (a) the original multi-spectral image with a 2.8m resolution which is used as the true-image to compare between the fusion result.

The results of the different fusion methods are shown in Fig. 7 (b)-(f). Visually, the MGF method obtains MS images with high-spatial-resolution, but the obvious color distortions occur in the playground. Spectral distortions from HCS method (i.e. shown in Fig. 7(c)) are clearly shown as color changes in the high reflection areas, such as, the green areas.

**C. Experiments with GEO-2 Data**

The results of the different fusion methods are shown in Fig. 6(b)-(f). We cut a sub-image, which is located at the right-bottom corner of the fused image, in order to demonstrate the fusion results more clearly. Visually, spectral distortions from MGF method (i.e. shown in Fig. 6(b)) are obvious as color changes severely in the red corner. The fused result obtained by the HCS method suffers from serious spectral distortion when compared with all the other methods. The AIHS method has improved, but the effect is not significant. To sum up, our method outperforms the spectral and the spatial resolution. The quantitative assessment indices results are listed in Table I, where the best results of every quality indices are clearly noticeable by font **bold**. Here, it can be clearly observed that our method performs the superior results in the six evaluation indexes when compared to the other methods.

**Fig. 5 Performance of our fusion method with respect to different guided filter stages.**

**Fig. 6 (a) Original Multi-spectral image. (b) MGF approach. (c) HCS approach. (d) AIHS approach. (e) First stage proposed method. (f) Second stage proposed method.**
The AIHS method has improved, but, it changes the color of some parts of the fused image when compared with the reference-image such as the playground. Fig. 7(e)-(f) verify that the proposed method results are more similar to the reference-image than the results of the other methods. The numerical experiments of the fused methods in Fig. 7 are listed in Table II. From Table II, it can be clearly observed that our fusion results have the best values for most of the quality indices.

### D. Experiments with GF-2 Data

In this part, we study and analyze the experimental-results on GF-2 data set. The GF-2 satellite captures 4m multi-spectral images, i.e., red, green, blue, NIR images, and 1m panchromatic image. In this experiment, the size of the panchromatic and the multi-spectral images are $1024 \times 1024$ and $256 \times 256$, respectively. As in the IKONOS-2 experiment, we spatially degrade the GF-2 data, then we utilize a simulated GF-2 LR multi-spectral image with the size of $64 \times 64$ and a corresponding panchromatic image with the size of $256 \times 256$. The results for GF-2 are presented in Fig. 8. The original-MS image with a 4m resolution is used as the true-image, in order to compare it with the fusion result is shown in Fig. 8(a).

The fusion results are compared with the original multi-spectral image, in order to obtain the desired objective measures, as shown in Fig. 8(b)-(f). By comparing the fusion results with the reference multi-spectral image visually, it can be observed that the fusion result of the HCS method suffers from a high-spectral-distortion. However, the spatial details in some areas are lost. The result obtained from the MGF method suffers from spectral distortion while losing some spatial details. As shown in Fig. 8(d), AIHS performs the fused result with well preserved spectral and spatial details. To sum up, our proposed method results is a much better result and a more similar image to the reference-image compared with the other methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>CC</th>
<th>RMSE</th>
<th>ERGAS</th>
<th>SAM</th>
<th>UIQI</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGF</td>
<td>0.9396</td>
<td>0.9168</td>
<td>0.9168</td>
<td>0.9422</td>
<td>0.9366</td>
<td>0.9429</td>
</tr>
<tr>
<td>AIHS</td>
<td>0.7502</td>
<td>0.7223</td>
<td>0.7368</td>
<td>0.7570</td>
<td>0.7558</td>
<td>0.7661</td>
</tr>
<tr>
<td>Proposed 1st stage</td>
<td>0.9495</td>
<td>19.9996</td>
<td>2.9292</td>
<td>5.7394</td>
<td>0.9465</td>
<td>0.7728</td>
</tr>
<tr>
<td>Proposed 2nd stage</td>
<td>0.9476</td>
<td>20.4610</td>
<td>2.9958</td>
<td>5.7643</td>
<td>0.9449</td>
<td>0.7731</td>
</tr>
</tbody>
</table>

**Table I**

**RESULTS COMPARISON OF OUR PROPOSED METHOD WITH SOME EXISTING APPROACHES ON IKONOS-2 DATA-SET**

---

Fig. 7 (a) Original Multi-spectral image. (b) MGF approach. (c) HCS approach. (d) AIHS approach. (e) First stage proposed method. (f) Second stage proposed method.
Fig. 8 (a) Original Multi-spectral image. (b) MGF approach. (c) HCS approach. (d) AIHS approach. (e) First stage proposed method. (f) Second stage proposed method.

Table III reports the quantitative assessments indexes calculated for GF-2 data set. From Table III, it can be clearly observed that our method outperforms all the other methods.

V. CONCLUSION

In this paper, a novel image fusion method based on an IHS injection model and a multi-stage guided filter is presented. In the first stage of the guided filter we used the $I$ component, which is acquired from the up-sampled multi-spectral image as the input image and the panchromatic image as the guidance-image, but in the second stage we used the panchromatic image as the guidance-image and the result of the first stage as the guidance-image, then the total semantic details map, which are extracted by the multi-stage guided filter and were injected into the up-sampled MS-image. Our contribution here is threefold. Firstly, the $I$ component is acquired from the up-sampled MS-image. Secondly, a multi-stage guided filter strategy was used in filtering the panchromatic image, in order to acquire more detail information. Third, the total semantic details map is injected into every band of the up-sampled multi-spectral image to acquire the fusion result, then, it makes an iterative back projection (IBP) for each band. Our method is tested on three degraded data-sets i.e., IKONOS-2, GEO-2, and GF-2, and compared with three existing methods. The empirical-results show that our fusion method is more robust in an image-registration among of the other methods. Furthermore, the proposed method is highly-effective among of the other methods.

**REFERENCES**


