Blind Fusion of Hyperspectral Multispectral Images Based on Matrix Factorization

Jian Long and Yuanxi Peng *

State Key Laboratory of High Performance Computing, College of Computer Science and Technology, National University of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China; jian_long@nudt.edu.cn
* Correspondence: pyx@nudt.edu.cn

Abstract: The fusion of low spatial resolution hyperspectral images and high spatial resolution multispectral images in the same scenario is important for the super-resolution of hyperspectral images. The spectral response function (SRF) and the point spread function (PSF) are two crucial prior pieces of information in fusion, and most of the current algorithms need to provide these two preliminary pieces of information in advance, even for semi-blind fusion algorithms at least the SRF. This causes limitations in the application of fusion algorithms. This paper aims to solve the dependence of the fusion method on the point spread function and proposes a method to estimate the spectral response function from the images involved in the fusion to achieve blind fusion. We conducted experiments on simulated datasets Pavia University, CAVE, and the remote sensing images acquired by two spectral cameras, Sentinel 2 and Hyperion. The experimental results show that our proposed SRF estimation method can improve the PSNR value by 5 dB on average compared with other state-of-the-art SRF estimation results. The proposed blind fusion method can improve the PSNR value of fusion results by 3–15 dB compared with other blind fusion methods.

Keywords: hyperspectral imaging super-resolution; image fusion; matrix factorization

1. Introduction

Hyperspectral images contain mapping information of the same scene in different wavelength bands, which means hyperspectral images have important applications in agriculture [1], medicine [2], geology [3], remote sensing [4], image processing [5,6], and other fields. The effectiveness of the application of a hyperspectral image is usually closely related to its spatial resolution. However, to ensure that each spectrum of the hyperspectral camera sensor has a sufficient number of photons, the hyperspectral camera needs to increase its viewing range, which causes the spatial resolution of the hyperspectral image to become lower. Compared with hyperspectral images, multispectral images have a wider spectral band and a more significant number of photons obtained per spectral band. Thus, multispectral images have a higher spatial resolution. A popular method to obtain high spatial resolution hyperspectral images (HR-HSI) is to fuse high spatial resolution multispectral images (HR-MSI) with low spatial resolution hyperspectral images (LR-HSI).

In hyperspectral image fusion, two important a priori pieces of information are usually involved, the point spread function (PSF): a fuzzy matrix used to model the spatial blur from the target HR-HSI to the LR-HSI; and the spectral response function (SRF): a spectral downsampling matrix for modeling the target HR-HSI to the HR-MSI. Many fusion methods [7–10] in the current study require the prior assumption that both SRF and PSF are known (non-blind fusion). However, PSF and SRF are not always known in practical fusion, which can lead to limitations in applying such methods. In further research, methods [11,12] that require only the assumption that SRF is known to complete the fusion were proposed (semi-blind fusion). However, in the absence of SRF, this method cannot be used either. To address the limitations of these fusion methods, direct estimation
of SRF and PSF from LR-HSI and HR-HSI has been proposed in some literature [13,14]. However, these SRF estimation methods have significant limitations and can only achieve good results in certain situations. To address this limitation, a new and more generalized SRF estimation method is proposed in this paper. Furthermore, to achieve blind fusion (no prior assumption that SRF and PSF are known in fusion), this paper develops a blind fusion algorithm without adjusting parameters.

In estimating SRF, we first perform an identical strong spatial blurring for LR-HSI and HR-MSI, according to the approach in [13] to counteract the effect from the true PSF. Furthermore, to solve the drawback that the HR-MSI and LR-HSI spectral bands need to be fixed in advance for the relationship when estimating SRF in [13]. Inspired by [14], we then build a quadratic programming problem to solve SRF based on the equation relationship between HR-MSI and LR-HSI after fuzzification. The SRF obtained at this point still has an error. To further optimize the SRF, we use the obtained SRF as an initial value and perform a matrix iterative multiplicative optimization operation on this initial value to obtain an accurate SRF estimate. When optimizing SRF using the matrix multiplication iterative process, we set an iterative optimization termination condition based on the equation relationship between HR-MSI and LR-HSI and prevent overfitting in the iterative process by the termination condition. Finally, we apply the estimated SRF to our proposed semi-blind fusion algorithm, which achieves a blind fusion of hyperspectral and multispectral images.

The schematic diagram of the fusion is shown in Figure 1, and the main innovations of this paper are as follows.

- An accurate SRF estimation method is proposed based on a quadratic programming model and matrix multiplication iterative solutions to achieve blind hyperspectral, multispectral image fusion.
- The estimated SRF is applied to our proposed semi-blind fusion, and the proposed blind fusion method is used to fuse real remote sensing images obtained by two satellite-based spectral cameras, sentinel 2 and Hyperion. The subjective results of the fusion demonstrate the superiority of our proposed method.
- An automatic parameter optimization method is proposed to automatically adjust the parameters during the fusion process, thus reducing the dependence of the fusion quality on the parameters.

The rest of this paper is organized as follows. Section 2 gives an introduction to related works. Section 3 introduces the fusion model based on matrix factorization. Section 4 details the proposed method. Section 5 gives the experimental settings. The results and analysis are presented in Section 6, and the conclusion is given in Section 7.
2. Related Works

One of the first methods to achieve blind fusion was the intensity-chromaticity-saturation (IHS) method [15], used to achieve the fusion of high spatial resolution panchromatic images (Pan) and low spatial resolution triple-frequency multispectral images. In the fusion, firstly, the IHS transform is applied to the three-band multispectral image, and then the Pan replacement I component after histogram alignment, and then the inverse IHS transform is applied to the replacement component to obtain the three-band multispectral image with high spatial resolution. However, since the histogram-matched Pan and intensity component I do not usually have the same radiometry, the spectral distortion occurs when using the color composition shows the fusion result. The reason for this spectral distortion is since the spectral response of the three-band multispectral image may differ from the spectrum of Pan, causing the appearance of local spatial details and spatially varying radiation shifts. A generalized IHS transform is proposed in the literature [16] to cope with the case where more than three bands are involved in the fusion when the
intensity component I is a weighted sum of the values of the individual bands, where the spectral response of Pan determines the weights. In some methods, these weights are determined directly as a definite value \[16,17\], which tends to cause the fused image to contain more spectral distortions, especially when the spectral bands of the multispectral image do not entirely cover the spectrum bands of Pan.

In a further study, the literature \[18\] proposed the principal component analysis (PCA) method, the same IHS similar, using the histogram-matched Pan image to replace the first principal component (PC1); in general, the fusion effect obtained using the PCA method is better than the IHS method. Another similar method is proposed in the literature \[19\]: the Gram-Schmidt (GS) spectral sharpening, which was used in the Environment for Image Visualization (ENVI) program package. In the GS method, for the conversion process of multispectral images, each spectral band coefficient is chosen as a fixed coefficient, leading to distortion of the fused spectra. In order to overcome the effects caused by fixed coefficients, in the literature \[20\], a multivariate regression method is proposed, which provides a more accurate estimation of the coefficients. It proposes the GSA and GIHSA methods based on GS and GIHS. This method can estimate the values of the coefficients in the fusion process more accurately, which can effectively improve the quality of the fusion and reduce the spectral distortion in the fusion. For fusing HR-MSI and LR-HSI, where each band of HR-MSI is considered as a Pan image, and the spectral bands in LR-HSI are divided into several band groups according to their correlation with the bands in HR-MSI, each band group with the corresponding Pan image are then fused using GSA or GIHSA. However, even with the multivariate regression technique, the GSA and GIHS still suffer from spectral distortion, and the image fusion quality still needs to be improved.

Besides, methods to achieve blind fusion also include smoothed filter-based intensity modulation (SFIMHS) \[21\], and Bayesian methods based on maximum a posteriori probability estimation stochastic mixture model (MAPSMM) \[22\]. SFIMHS is a development of pixel block intensity modulation (PBIM), a method of fusion based on a simplified solar radiation and ground reflection model. This approach can also be applied to the fusion of LR-HSI and HR-MSI after a generalized extension. Similar to the IHS technique, this method adjusts the retention of spatial detail between low spatial resolution multispectral images and high spatial resolution Pan images by using coefficients between the high spatial resolution images and its low-pass smoothed filtered images. However, this method also suffers from spectral distortion, and the fusion quality still needs further improvement. The MAPSMM method is a blind fusion method that optimizes in the principal component subspace. This method obtains LR-HSI’s mean spectrum, covariance matrix, and abundance map of each end element. It builds a MAP target function based on SMM statistics to optimize the HR-HSI data relative to the input image. Since this is a fusion model based on statistical estimation, the fusion effect of this method also needs further improvement.

The matrix factorization-based fusion method, which assumes the target HR-HSI as the product of the spectral basis matrix and the coefficient matrix, has emerged in the subsequent studies of the fusion method, and the fused HR-HSI is obtained by estimating the spectral basis and coefficient matrix. The literature \[23\] obtained high-quality fusion results by alternating non-negative matrix factorization (NMF) for HR-MSI and LR-HSI. In a subsequent study in \[14\], after aligning the HR-MSI and LR-HSI involved in the fusion, the HR-MSI was first resampled to the same resolution as the LR-HSI, followed by estimating the SRF from the obtained spectrally downsampled images of the LR-MSI and LR-HSI. After the SRF is estimated, the PSF is obtained by solving a minimum quadratic formulation problem with constraints. The PSF matrix finally obtained in the literature \[14\] is a Gaussian blur filter. For each multispectral band, the parameters of Gaussian blur are obtained by maximizing the maximum correlation coefficient between the Sobel image of the image LR-MSI and LR-HSI spectrum down-sampling (The Sobel operation is a sharpness common edge detection algorithm used to compute image sharpness \[24\]). However, this method fixes the PSF as a Gaussian blur in advance, which makes the
estimation of PSF and SRF of this method limited, and the fusion effect will be poor for the non-Gaussian blur case.

The method of vector total variance based on subspace regularization is proposed in the literature [13] to solve the fusion problem, which is also a matrix factorization method. Ref. [13] also proposed PSF and SRF estimation. In [13], a strong spatial blur is first performed simultaneously for LR-HSI and HR-MSI, which is used to weaken the influence from the truth PSF matrix, thus facilitating a more accurate estimation of the SRF. The estimated SRF is then brought into a quadratic matrix equation to find a more accurate estimate of the PSF. Finally, the obtained PSF and SRF are used to calculate the estimates of the spectral basis and coefficient matrices, thus enabling blind fusion.

Among some recent studies, semi-blind fusion methods have arisen, i.e., methods that only need to provide SRF to achieve hyperspectral, multispectral image fusion. In the literature [11], a semi-blind fusion method based on coupled tensor canonical polyadic decomposition (CPD) is proposed. It factorizes the HR-HSI into a product of three-dimensional factor matrices and achieves fusion by estimating the factor matrices of the three dimensions. However, this approach requires the assumption that HR-HSI has the same rank in the three dimensions, and the three ranks are usually different, which often leads to a compromised effect of fusion. The semi-blind fusion method based on non-local sparse tensor factorization is also proposed in the literature (NLSTF-SMBF) [12]. NLSTF-SMBF is a fusion method based on tucker decomposition and adds prior information of non-local clustering to the fusion, which improves the quality of fusion. However, this method usually involves many variables and is computationally intensive, taking a lot of computation time.

Table 1 summarizes the strengths and weaknesses of the above-mentioned blind fusion methods and our proposed method’s strengths and possible improvements.

Table 1. Summary of the strengths and weaknesses of related methods.

| Method | Strengths | Weaknesses/Improvement |
|--------|-----------|------------------------|
| Component substitution based methods: IHS [15], PCA [18], GIHS [20], GSA [20] | Less time-consuming | Spectral distortion, lower fusion quality |
| Smoothed filter-based intensity modulation method: SFIMHS [21] | Least time consuming | Spectral distortion, lower fusion quality |
| Statistical estimation-based fusion method: MAPSMM [22] | Robust fusion quality on different datasets | The longest fusion time-consuming and the worst fusion quality |
| Matrix factorization-based fusion methods: Hysure [13], CNMF [14,23] | Included SRF and PSF estimation can be extended to non-blind fusion algorithms | More time-consuming, constrained fusion quality |
| The proposed method | Higher fusion quality, faster fusion speed, better SRF estimation | Add a priori information to further improve fusion quality |

3. Fusion Model

We use matrices to describe the images involved in the paper, recording the target HR-HSI as $Z \in \mathbb{R}^{L \times P}$, denoting $L$ spectral bands and $P$ pixel points; HR-MSI as $M \in \mathbb{R}^{L' \times P}$ with $L'$ spectral bands and $P$ pixel points; and LR-HSI as $H \in \mathbb{R}^{L' \times p}$ with $L'$ spectral bands and $p$ pixel points.

Generally, LR-HSI can be regarded as a spatially blurred downsampled image of the target HR-HSI, i.e., LR-HSI is obtained by doing a spatial blur HR-HSI followed by spatial downsampling. HR-MSI can be regarded as a spectral downsampling matrix of the target HR-HSI. The following equation can express it.

$$H = ZBS$$  \hspace{1cm} (1)

$$M = RZ$$  \hspace{1cm} (2)

where $B \in \mathbb{R}^{P \times P}$ represents the spatial fuzzy matrix and can be used to model the role of PSF, and $S \in \mathbb{R}^{P \times p}$ represents the spatial downsampling matrix, and $R \in \mathbb{R}^{L \times L'}$ represents the spectral downsampling matrix, which can be used to simulate the role of SRF. In blind
fusion, both PSF and SRF are unknown, i.e., the matrices B and R are unknown. In the matrix factorization-based model [8], the target HR-HSI is decomposed into the product of the spectral basis matrix and the coefficient matrix, i.e., \( Z = DC \) \(^{(3)}\)

where \( D \in \mathbb{R}^{L_s \times L_s} \) denotes the spectral basis with an \( L_s \)-dimensional subspace and \( C \in \mathbb{R}^{L_s \times P} \) denotes the coefficient matrix. Taking Equation (3) into Equations (1) and (2) yields the representation of the matrix factorization of LR-HSI and HR-MSI.

\( H = DCBS \) \(^{(4)}\)

\( M = RDC \) \(^{(5)}\)

4. Proposed Method

This section presents our proposed method. Our proposed method mainly includes two steps: estimation of SRF and blind fusion using SRF. When estimating SRF, we first perform a spatial down-sampling of HR-MSI, then perform the same strong spatial fuzzing processing with LR-HSI, and then estimate SRF through a quadratic programming mathematical model and multiplication iterative process. In blind fusion, the estimated SRF and HR-MSI are first used to obtain an approximate estimate of HR-HSI based on the Moore–Penrose inverse method proposed in our previous work [10]. Then, the spectral information part of the approximate estimate is replaced by the spectral information in LR-HSI using two SVDs. Finally, an exact HR-HSI estimate is obtained using a multiplicative iterative process.

4.1. Estimation of SRF

From Equations (1) and (2), we can obtain the following equations.

\( RH = MBS \) \(^{(6)}\)

In the Equation (6), there are two unknown matrices \( R \) and \( B \), so the SRF cannot be solved directly. To solve this problem, inspired by the method in [13], we multiply both sides of the equal sign by a stronger spatial fuzzy matrix \( Bm \). Since the \( Bm \) fuzzy strength is much larger than \( B \), the effect of \( B \) can be ignored. It is worth noting that we differ from the method in [13], in which a strong spatial fuzzy matrix \( Bm_1 \) is used to fuzz the MSI first, and then a fuzzy matrix \( Bm_2 \) with the same vigorous intensity as \( Bm_1 \), but with the same spatial dimension as the HSI is used to fuzz the HSI. Although \( Bm_1 \) and \( Bm_2 \) have the same fuzzy intensity, they still introduce errors due to the different dimensions of these two matrices. To solve this problem, we first do spatial downsampling of the MSI. After sampling, we obtain a multispectral image with the same spatial size as the HSI, after which we blur the sampled MSI and HSI using a blurring matrix \( Bm \) with the same intensity.

\( RHBm = MBSBm \) \(^{(7)}\)

In Equation (7), since the intensity of the fuzzy matrix \( Bm \) is much larger than \( B \), the effect of \( B \) can be ignored so that we can approximate the following equation.

\( MBSBm \approx MSBm \) \(^{(8)}\)

From this, we can obtain the following equation.

\( RHBm \approx MSBm \) \(^{(9)}\)
In Equation (9), there is only one unknown \( R \). For solving \( R \), we build a quadratic programming mathematical model, we solve \( R \) row by row, treating each row of \( R \) as an unknown vector, which yields the following equation.

\[
\frac{1}{2} R_i (HBm)(HBm)^T R_i^T - \frac{1}{2} (MSBm)_i (HBm)^T R_i^T \approx 0
\]

where \( R_i \) represents the i-th row vector of \( R \), \( (MSBm)_i \) represents the i-th row of the matrix after fuzzing by MSI downsampling, in Equation (10), \( (HBm)(HBm)^T \) is the Hamilton matrix, and the quadratic programming mathematical model can be obtained by variable substitution as follows.

\[
\min_x \frac{1}{2} x^T H_0 x + f^T x, \text{ s.t. } x \geq 0 \text{ where } \begin{cases} x = R_i^T \\ H_0 = (HBm)(HBm)^T \\ f = -\frac{1}{2} (HBm)(MSBm)_i^T \end{cases}
\]

We can obtain the initial solution \( R_1 \) of the spectral response function \( R \) by solving the quadratic programming problem of Equation (11). However, since there are still uncertainties in the quadratic programming, we need further precise values of \( R \). One further optimization equation can be obtained according to Equation (9).

\[
\min_R \|RHBM - MSBm\|_F^2
\]

This is a convex problem, and inspired by the multiplicative update rule in the literature [23], we solve the exact solution of \( R \) by the following multiplicative iterative optimization rule. \( R_1 \) is brought into the calculation as the initial value of \( R \) in the iteration, and the solution obtained after the iterative optimization is the exact solution of \( R \).

\[
R = R \times \left( (MSBm)(HBm)^T \right) / \left( R(HBm)(HBm)^T \right)
\]

In summary, the estimation process of SRF can be shown by Algorithm 1.

**Algorithm 1 Estimated SRF**

**Require:** HM

Perform spatial downsampling of the MSI matrix M.

Using the strong fuzzy matrix Bm fuzzy matrix H, MS.

for \( k = 1 : \text{size}(M,1) \) do

\( x = \text{solution of Equation (9)} \)

\( R_{1_k} = x^T \)

end for

\( R = R_1 \)

for \( k = 1 : K \) do

Update \( R \) by Equation (13)

end for

return \( R \)

**4.2. Blind Fusion Method**

After estimating the SRF, we devise a simple semi-blind algorithm to solve the fusion problem and thus blind fusion. The preliminary estimation matrix \( Z_1 \) of the target HR-HSI can be obtained based on our previous work [10].

\[
Z_1 = \text{Moore-Penrose}(R)M
\]

At this time, \( Z_1 \) mainly contains the spectral information and spatial information in \( M \), and lacks the spectral information in \( H \). In the literature [8,9], it was mentioned and
verified that LR-HSI and the target HR-HSI contain the same spectral basis, and SVD can obtain the spectral basis.

\[ H = U \Sigma V^T \]  
\[ Z = \hat{U} \hat{\Sigma} V^T \]  

The spectral basis of hyperspectral images is mentioned in the literature \[8\] in a lower dimension. By keeping the first \( q \) maximum singular values, the following Equation (17) can be obtained.

\[ H_0 = U_0 \Sigma_0 V_0^T \]  
\[ Z_0 = \hat{U}_0 \hat{\Sigma}_0 \hat{V}_0^T \]  

So, the spectral basis \( D \) can be represented as

\[ D = \hat{U}_0 = \hat{U}(;1 : q) \approx U_0 = U(;1 : q) \]  

The correlation between the singular values of LR-HSI and HR-HSI was verified by extensive experiments in our previous work \[9\], and further studies revealed that the singular values of HR-HSI are \( sf \) times the singular values of LR-HSI. \( sf \) represents the spatial downsampling factor from HR-HSI to LR-HSI. Inspired by these works, we obtain a further estimate of the target HR-HSI \( Z_2 \), by performing an SVD decomposition of \( Z_1 \) and \( H \) simultaneously and replacing the spectral basis in \( Z_1 \) with the spectral basis obtained from the decomposition in \( H \).

\[ [U_1, \Sigma_1, V_1] = \text{svds}(Z_1, q) \]  
\[ [U_2, \Sigma_2, V_2] = \text{svds}(H, q) \]  
\[ Z_2 = U_2(sf \times \Sigma_2)V_1^T \]  

where \( q \) denotes the dimensionality of the spectral basis, \( sf \) denotes the downsampling factor, the value of \( q \) affects the quality of fusion, and our extensive experiments have shown that the value of \( q \) is between 1 and 6. We set the MSI obtained by multiplying the fused HR-HSI and SRF with the similarity of the MSI involved in the fusion as a criterion to automatically traverse the best \( q \)-value from 1 to 6, thus realizing the automatic tuning of our fusion method.

After obtaining \( Z_2 \), we then bring \( Z_2 \) as an initial value into the multiplicative iterative optimization equation of Equation (23) based on our previous work to find the exact solution \( Z \) of the target HR-HSI.

\[ Z = Z_2 \times (R^T M) . / (R^T R Z) \]  

In summary, our proposed blind fusion algorithm with automatic parameter tuning is shown below Algorithm 2.
**Algorithm 2 Blind Fusion**

**Require:** $H, M$

- Get $R$ by Algorithm 1
- Get $Z_1$ by Equation (14)
- Set the `minvalue` and `maxq` to an initial value

```latex
\text{for } q=1:6 \text{ do}
\begin{align*}
\begin{bmatrix} U_1, \Sigma_1, V_1 \end{bmatrix} &= \text{svds}(Z_1, q) \\
\begin{bmatrix} U_2, \Sigma_2, V_2 \end{bmatrix} &= \text{svds}(H, q) \\
Z_2 &= U_2(sf \times \Sigma_1)V_1^T \\
\text{min}_1 &= RZ_2 - M \\
\text{if } \text{min}_1 < \text{minvalue} \text{ then} \\
\text{minvalue} &= \text{min}_1 \\
\text{maxq} &= q
\end{align*}
\text{end if}
\text{end for}

$q = \text{maxq}$

$$\begin{bmatrix} U_1, \Sigma_1, V_1 \end{bmatrix} = \text{svds}(Z_1, q)$$
$$\begin{bmatrix} U_2, \Sigma_2, V_2 \end{bmatrix} = \text{svds}(H, q)$$
$$Z_2 = U_2(sf \times \Sigma_1)V_1^T$$
$$Z = Z_2$$
\text{for } k=1:K \text{ do}
  \begin{align*}
  \text{Update } Z \text{ by Equation (23)}
\end{align*}
\text{end for}
\text{return } Z
```

5. **Experiments**

To validate the effectiveness of our proposed method, we have done experiments on both remote sensing datasets ground datasets and added experimental comparisons on real remote sensing datasets. In addition, to demonstrate the effectiveness of our proposed method in all aspects, we also performed the fusion of simulated LR-HSI and HR-MSI generated in two different SRF and two different PSF cases.

5.1. **Dataset**

We used two simulated datasets, the remote sensing dataset Pavia University [25] and the ground dataset CAVE [26], where Pavia University is a remote sensing image acquired by the ROSIS sensor over Pavia University, the original Pavia University image contains 115 bands each band contains $610 \times 340$ pixels. After removing some absorbing bands and some undesirable pixels, the final size of the simulated HR-HSI we used in our experiments was $256 \times 256 \times 93$, and the spatial resolution of the Pavia University image was 1.3 m. The spectral bands ranged from 0.43 to 0.86 um. We used two different PSFs to obtain different LR-HSI size $64 \times 64 \times 93$ (one is a $7 \times 7$ (standard deviation 2) Gaussian blur other is a $3 \times 3$ average blur). We used two different SRFs to obtain different HR-MSIs, one SRF for the IKONOS class reflection spectral response filter [27] to obtain an HR-MSI size of $256 \times 256 \times 4$, and the other SRF we made using Gf-1-16 m multispectral camera response to obtain HR-MSI size of $256 \times 256 \times 3$.

The ground dataset CAVE [26] contains 32 high-quality indoor HSIs, the image resolution is $512 \times 512$, the spectral band range is 400–700 nm, and the spectral resolution is 10 nm. Each image contains 31 spectral bands, but the first two bands are blurred, so in the experiment, we use the size of the CAVE image as $512 \times 512 \times 29$. We also used the two PSFs mentioned above to generate the LR-HSI, down sampling every 8 pixels in the HR-HSI to obtain an LR-HSI size of $64 \times 64 \times 31$. As with Pavia University, two different SRFs (Use the first 29 bands) were used to generate the HR-MSI with sizes of $512 \times 512 \times 4$ and $512 \times 512 \times 3$, respectively.

Finally, we also conducted experiments on a real remote sensing dataset. The LR-HSI in the real remote sensing dataset is an image with 220 spectral bands, 30 m ground spatial...
resolution, and a 400–2500 nm spectral range acquired by the Hyperion sensor on the Earth Observation 1 satellite. After removing the absorbing bands in practical applications, 89 bands are retained, while we use a $100 \times 100 \times 89$ size area as the fused LR-HSI in the experiments of this paper. The HR-MSI in the real remote sensing dataset is an image containing 13 spectral bands with a spatial ground resolution of 10 m acquired by the Sentinel-2A satellite. In this paper, we select central four bands of size $300 \times 300 \times 4$ as the HR-MSI in fusion.

5.2. Compare Method and Parameter Setting

Firstly, we use two latest semi-blind fusion algorithms, Super-resolution TEnsor-REcOnstruction (STEREO) [11] (we use the semi-blind version) and non-local sparse tensor factorization for semi-blind fusion (NLSTF_SMBF) [12], to compare the effect of SRF estimated by our method with that estimated by [13,14]. For fair and reliable experimental results, the parameters of STEREO (https://sites.google.com/site/harikanats/ (accessed on 6 June 2021)) and NLSTF_SMBF (https://github.com/renweidian/NLSTF (accessed on 6 June 2021)) were kept constant in the experiments. The parameters of STEREO are set to $t_{\text{rank}} = 190, \text{maxit} = 45, \lambda = 210, \text{bs}_\text{iter} = 60$; and the parameters of NLSTF_SMBF are set to $W = 10, H = 10, S = 14, \lambda = 1e-6, K = 151, \lambda_1 = 1 \times 10^{-5}, \lambda_2 = 1 \times 10^{-5}, \lambda_3 = 1 \times 10^{-5}$.

Then, we compare our proposed method with five different blind fusion methods. These blind fusion methods include methods based on matrix factorization Hysure (https://github.com/alfaiate/HySure (accessed on 6 June 2021)) [13], CNMF [23]; Hysure and CNMF have their own proposed estimation of SRF and PSF. The method GSA based on component replacement [20]; the method SFIMHS based on a simplified model of solar radiation and land surface reflection [21]; the fusion model MAPSMM based on maximum a posteriori estimation [22]. To ensure the fairness of blind fusion, where the PSF and SRF in the Hysure method are derived from the self-contained PSF and SRF estimation methods in [13], and the CNMF also uses the self-contained PSF and SRF estimation methods in [14], our proposed method uses the SRF estimation methods we propose in this paper. The GSA, SFIMHS, and MAPSMM methods are among the fusion methods that do not require PSF and SRF. Our proposed method, GSA, and SFIMHS do not require setting parameters in terms of parameter setting. For the other methods, since only subjective results of fused images can be obtained in real dataset fusion and it is not possible to modify the parameters according to accurate, objective indicators, we keep the parameter settings of Hysure, CNMF, and MAPSMM in the original literature in order to ensure the realistic and reliable comparison experiments.

The experiments were implemented in MATLAB R2014a on localhost with 2.40 GHz Intel i7-7700 CPU and 8.0 GB DDR3.

5.3. Quantitative Metrics

In order to evaluate the effectiveness of our proposed method in all aspects, we use a combination of subjective and objective evaluation metrics. For the subjective metrics, we show the effectiveness of our proposed method through the visual presentation of the fused images. The objective indexes we used are peak signal-to-noise ratio (PSNR) [8], which can be used to evaluate the spatial reconstruction quality of each band, the larger the value, the better the fusion effect; root mean square error (RMSE) [8], which is a measure of the estimation error, the smaller the value, the better the fusion quality; relative dimensionless global error (ERGAS) [28], which reflects the overall quality of the fused image, the smaller the value, the better the fusion quality; spectral angle mapper (SAM) [27], which is used to evaluate the angle between the fused pixel and the real pixel in the whole spatial domain, the smaller the value, the better the fusion effect; structural similarity (SSIM) [29], which is used to evaluate the similarity of two images by three aspects: brightness, contrast, and structure, the closer the value to 1, the better the effect; fusion time, which is used to evaluate the speed of the fusion algorithm, the smaller the value, the faster the fusion.
6. Results and Analysis

We demonstrate the performance of our proposed method from three aspects:

- We will examine the superior performance of our proposed SRF estimation using two state-of-the-art semi-blind fusion methods.
- We will compare our proposed method with five blind fusion methods to demonstrate the advanced performance of our proposed blind fusion.
- We will visually compare the effect of our proposed blind fusion method with the five blind fusion methods on the real dataset to highlight the practicality of our proposed method.

6.1. Advancement of the Proposed SRF Estimation Algorithm

First, we label the two SRFs used in this paper as SRF1 and SRF2, where SRF1 denotes the IKONOS class reflection spectral response filter [27] and SRF2 denotes the spectral response function produced using the Gf-1-16m multispectral camera response. Note that since only 29 spectra were used in the CAVE dataset and 93 spectra were used in the Pavia University dataset, in the CAVE dataset, SRF1 and SRF2 are new SRFs generated from the first 29 bands of both SRFs. We also labeled the two PSFs used in the experiment as PSF1 and PSF2, where PSF1 denotes a $7 \times 7$ (standard deviation of 2) Gaussian blur and PSF2 denotes a $3 \times 3$ average blur. In order to comprehensively verify the superior performance of our proposed SRF estimation method, we use two semi-blind algorithms, STEREO and NLSTF, using our estimated SRF, literature [13] estimated SRF (Recorded as R1 in the table), literature [14] estimated SRF (Recorded as R2 in the table), and actual used SRF to complete fusion work. The objective fusion metrics of the fusion result under different SRF show the superior performance of our proposed estimation method.

Table 2 shows the results of our proposed method on the Pavia University dataset. The first column of the table indicates the types of PSF and SRF simulated to generate LR-HSI and HR-MSI in the experiment. To make the experimental results more adequate and reliable, we performed experiments under different PSFs and SRFs. Since our algorithm mainly targets SRF estimation and the semi-blind fusion is robust to PSFs, to reduce the redundancy of the experimental results, we only perform a set of comparison experiments when SRF2 is the same. The experimental results show that our proposed estimation method can estimate the SRF more accurately.

In Table 2, in addition to the fusion results using the real SRF, the best results of the other three SRF fusions are expressed in **bold**. As can be seen from Table 2, the results of SRF estimated using our proposed method are not much different from the results of using real SRF. Except that the SRF estimated by [13] when the STEREO method is in SRF1 and PSF1 is slightly better than ours, in other cases, the SRF estimated by our method has the best results. The estimated SRF using our proposed method can increase the PSNR value of the fusion results of the two semi-blind fusion methods by nearly 1dB, which is only less than 1dB less than using the real SRF.

Table 3 shows the fusion effect of using different SRFs on the CAVE dataset, which intuitively demonstrates the superior performance of our method. The fusion results of the SRF estimated using our method are similar to those of the real SRF, and the difference between the fused PSNR values is only within 1 dB. However, the fusion results of the SRFs estimated using the methods in [13,14] are not satisfactory, and the difference between the fused PSNR and the real SRF is even up to 15 dB. Combining Tables 2 and 3, the SRF estimated by [13] works well on the Pavia University dataset, while the SRF estimated by [14] works well on the CAVE dataset. The main reason for this result is that the SRF estimation method in [13] needs to force the hyperspectral band to correspond to the multispectral band, which requires the SRF value to be staggered in each row. In the CAVE dataset, because we only use the spectral response of the first 29 bands to make SRF, the real SRF used in the CAVE dataset has a value on each row, so the effect of [13] becomes poor. On the contrary, the estimation method in [14] is more suitable for the situation where SRF has a value in each row. However, in the Pavia University dataset, because the value
in the spectral response function is staggered on each row, the SRF estimated by [13] is better than [14]. While our proposed method integrates various distributions of the SRF values, the proposed method achieves the best fusion results on both datasets.

Table 2. Fusion results of objective metrics after two semi-blind methods using different SRFs on the Pavia University dataset.

| SRF & PSF Method        | SRF | PSNR | RMSE | ERGAS | SAM  | SSIM |
|-------------------------|-----|------|------|-------|------|------|
| NLSTF-SMBF [12]         | R1  [13] | 41.566 | 2.25 | 1.285 | 2.271 | 0.987 |
|                         | R2  [14] | 32.29 | 6.991 | 3.62 | 5.827 | 0.969 |
|                         | Ours | 41.836 | 2.175 | 1.262 | 2.176 | 0.987 |
|                         | SRF1 | 42.8 | 1.995 | 1.145 | 2.051 | 0.988 |
| STEREO [11]             | R1  [13] | 37.31 | 3.578 | 2.093 | 3.56 | 0.958 |
|                         | R2  [14] | 32.166 | 6.757 | 3.644 | 5.562 | 0.929 |
|                         | Ours | 37.223 | 3.622 | 2.129 | 3.643 | 0.956 |
|                         | SRF1 | 37.564 | 3.482 | 2.035 | 3.534 | 0.958 |
| NLSTF-SMBF [12]         | R1  [13] | 40.605 | 2.646 | 1.598 | 2.479 | 0.985 |
|                         | R2  [14] | 34.791 | 5.293 | 2.716 | 3.746 | 0.981 |
|                         | Ours | 41.522 | 2.342 | 1.423 | 2.289 | 0.987 |
|                         | SRF1 | 42.44 | 2.196 | 1.336 | 2.214 | 0.987 |
| STEREO [11]             | R1  [13] | 36.813 | 3.932 | 2.304 | 3.79 | 0.954 |
|                         | R2  [14] | 32.882 | 6.405 | 3.433 | 5.222 | 0.941 |
|                         | Ours | 37.225 | 3.734 | 2.265 | 3.797 | 0.955 |
|                         | SRF1 | 37.501 | 3.636 | 2.187 | 3.733 | 0.956 |
| NLSTF-SMBF [12]         | R1  [13] | 40.991 | 2.537 | 1.516 | 2.416 | 0.986 |
|                         | R2  [14] | 26.767 | 15.166 | 6.882 | 6.033 | 0.949 |
|                         | Ours | 41.311 | 2.396 | 1.478 | 2.268 | 0.987 |
|                         | SRF1 | 42.721 | 2.147 | 1.299 | 2.197 | 0.988 |
| STEREO [11]             | R1  [13] | 36.884 | 3.91 | 2.304 | 3.78 | 0.954 |
|                         | R2  [14] | 25.716 | 16.854 | 5.941 | 24.748 | 0.707 |
|                         | Ours | 37.109 | 3.759 | 2.3 | 3.823 | 0.955 |
|                         | SRF1 | 37.517 | 3.62 | 2.18 | 3.707 | 0.957 |

Table 3. Fusion results of objective metrics after two semi-blind methods using different SRFs on the CAVE dataset.

| SRF & PSF Method        | SRF | PSNR | RMSE | ERGAS | SAM  | SSIM |
|-------------------------|-----|------|------|-------|------|------|
| NLSTF-SMBF [12]         | R1  [13] | 25.061 | 15.307 | 5.588 | 29.181 | 0.813 |
|                         | R2  [14] | 33.202 | 5.897 | 2.358 | 8.061 | 0.944 |
|                         | Ours | 39.526 | 3.160 | 1.299 | 6.171 | 0.963 |
|                         | SRF1 | 40.644 | 2.638 | 1.069 | 5.569 | 0.972 |
| STEREO [11]             | R1  [13] | 26.122 | 13.515 | 4.903 | 27.599 | 0.807 |
|                         | R2  [14] | 32.687 | 6.092 | 2.395 | 9.433 | 0.907 |
|                         | Ours | 36.692 | 3.966 | 1.596 | 7.900 | 0.934 |
|                         | SRF1 | 37.213 | 3.664 | 1.456 | 7.557 | 0.943 |
| NLSTF-SMBF [12]         | R1  [13] | 27.123 | 11.73 | 4.382 | 21.584 | 0.824 |
|                         | R2  [14] | 33.233 | 5.681 | 2.18 | 10.397 | 0.94 |
|                         | Ours | 38.578 | 3.618 | 1.469 | 8.257 | 0.957 |
|                         | SRF2 | 39.45 | 3.202 | 1.281 | 7.954 | 0.965 |
| STEREO [11]             | R1  [13] | 24.93 | 16.854 | 5.941 | 24.748 | 0.707 |
|                         | R2  [14] | 32.471 | 6.26 | 2.457 | 11.122 | 0.913 |
|                         | Ours | 36.166 | 4.386 | 1.769 | 9.482 | 0.923 |
|                         | SRF2 | 36.604 | 4.066 | 1.616 | 9.134 | 0.935 |
### Table 3. Cont.

| SRF & PSF | Method          | SRF     | PSNR   | RMSE  | ERGAS  | SAM     | SSIM  |
|-----------|-----------------|---------|--------|-------|--------|---------|-------|
| NLSTF-SMBF [12] | R1 [13]        | 26.864  | 12.027 | 4.491 | 21.869 | 0.82    |       |
|            | R2 [14]        | 31.486  | 7.139  | 2.786 | 13.62  | 0.912   |       |
|            | **Ours**       | 38.479  | 3.702  | 1.507 | **8.326** | **0.961** |       |
| SRF2 PSF2  | SRF2           | 38.818  | 3.54   | 1.425 | 7.936  | 0.965   |       |
| STEREO [11] | R1 [13]        | 24.372  | 18.354 | 6.418 | 25.999 | 0.682   |       |
|            | R2 [14]        | 30.361  | 7.926  | 3.105 | 14.658 | 0.821   |       |
|            | **Ours**       | 36.384  | 4.226  | 1.696 | **9.235** | **0.928** |       |
|            | SRF2           | 36.629  | 4.047  | 1.608 | 9.008  | 0.935   |       |

### 6.2. Advancement of the Proposed Blind Fusion Algorithm

To comprehensively evaluate the superior performance of our blind fusion algorithm, we compare the subjective and objective fusion results of our proposed method and the other five blind fusion methods when simulating LR-HSI and HR-MSI obtained by different SRF and PSF.

Table 4 shows the fusion results of our proposed method and comparison method on the Pavia University dataset. It can be seen that our method can consistently achieve the best results on the five fusion objective evaluation indicators under different SRF and PSF. In the best case, the PSNR value of the image after the fusion of our method can be nearly 9 dB higher than the PSNR value obtained by the other five methods and at least 3 dB higher. It fully embodies the advanced fusion performance of our proposed method on remote sensing datasets. In terms of fusion time, in addition to the GSA and SFIMHS methods, the fusion time of our method is also the least, and because GSA and SFIMHS only involve simple image transformation operations, the fusion result is much worse than our method. Figure 2 shows the fusion results of the comparison methods on the Pavia University dataset. To show the effect of various fusion methods more visually, we enlarge the red box part of the figure to show the top left corner of the figure. The MAPSMM method and the SFIMHS method obtain the most blurred fused images, which is due to the fact that these two methods use only simple image transformations to obtain fused images, resulting in poor image detail retention. Our proposed method obtains better fusion results at different SRF and PSF. The images have better detail retention ability.

Table 5 shows the fusion results of our proposed method and the comparison method on the oil_painting images in the CAVE dataset. Our proposed method achieves the best results on four objective evaluation metrics where the PSNR value is 4 dB higher than the best of the other methods. However, due to the small number of spectral bands on the CAVE dataset and the SRF used has values in each row, the similarity between pixels of the fused results is relatively significant, so our fused results will have a large SAM. Collectively, our proposed method obtains better fusion results than other methods, and in terms of time consumption, our proposed method is much less time-consuming than other methods except for two methods involving only simple image transformations, GSA and SFIMHS. Figure 3 shows the subjective fusion results of the compared methods on the oil_painting images in the CAVE dataset, and again we enlarge the part in the red box to show it in the upper left corner. From the figure, it can be visualized that the stronger detail retention of our proposed method.
Table 4. Objective evaluation metrics for the fusion results of the six blind fusion methods in the paper on the Pavia University dataset.

| Pavia University | Method | SRF & PSF | PSNR | RMSE | ERGAS | SAM | SSIM | Time |
|------------------|--------|-----------|------|------|-------|-----|------|------|
| SRF1 PSF1        | Hysure [13] | 35.581 | 4.501 | 2.637 | 3.852 | 0.972 | 64.431 |
|                  | CNMF [14,23] | 31.307 | 7.019 | 4.181 | 3.855 | 0.934 | 25.953 |
|                  | GSA [20] | 34.671 | 4.966 | 3.049 | 4.087 | 0.962 | 0.839 |
|                  | MAPSMM [22] | 27.508 | 10.855 | 6.381 | 4.910 | 0.854 | 92.327 |
|                  | SFIMHS [21] | 28.461 | 9.736 | 5.771 | 4.129 | 0.892 | 0.234 |
|                  | Ours | 38.632 | 3.593 | 1.847 | 2.940 | 0.982 | 7.649 |
| SRF2 PSF1        | Hysure [13] | 33.53  | 6.585 | 3.522 | 4.983 | 0.956 | 69.896 |
|                  | CNMF [14,23] | 32.411 | 6.236 | 3.465 | 3.304 | 0.947 | 25.823 |
|                  | GSA [20] | 34.98  | 4.813 | 2.931 | 4.004 | 0.966 | 0.865 |
|                  | MAPSMM [22] | 27.214 | 10.631 | 6.215 | 4.824 | 0.861 | 87.706 |
|                  | SFIMHS [21] | 28.494 | 9.711 | 5.762 | 4.183 | 0.893 | 0.234 |
|                  | Ours | 40.032 | 3.34  | 1.72  | 2.890 | 0.985 | 7.639 |
| SRF2 PSF2        | Hysure [13] | 31.579 | 8.033 | 4.268 | 5.603 | 0.921 | 68.318 |
|                  | CNMF [14,23] | 31.556 | 7.281 | 3.786 | 4.077 | 0.920 | 25.434 |
|                  | GSA [20] | 32.474 | 6.286 | 3.537 | 4.967 | 0.955 | 0.847 |
|                  | MAPSMM [22] | 26.182 | 12.735 | 7.329 | 5.649 | 0.810 | 91.430 |
|                  | SFIMHS [21] | 26.394 | 12.365 | 7.256 | 5.001 | 0.840 | 0.235 |
|                  | Ours | 40.872 | 2.665 | 1.555 | 2.590 | 0.955 | 9.601 |

Table 5. Objective evaluation metrics for the fusion results of the six blind fusion methods in the paper on the CAVE dataset.

| CAVE | Method | SRF & PSF | PSNR | RMSE | ERGAS | SAM | SSIM | Time |
|------|--------|-----------|------|------|-------|-----|------|------|
| SRF1 PSF1 | Hysure [13] | 28.873 | 9.859 | 3.591 | 19.348 | 0.852 | 271.399 |
|          | CNMF [14,23] | 30.151 | 8.347 | 3.311 | 7.279 | 0.861 | 60.043 |
|          | GSA [20] | 33.646 | 5.646 | 2.262 | 10.210 | 0.915 | 1.305 |
|          | MAPSMM [22] | 27.069 | 11.638 | 4.563 | 7.949 | 0.849 | 200.775 |
|          | SFIMHS [21] | 26.309 | 12.951 | 5.136 | 6.834 | 0.883 | 0.434 |
|          | Ours | 37.889 | 3.925 | 1.644 | 11.791 | 0.943 | 10.98 |
| SRF2 PSF1 | Hysure [13] | 30.872 | 7.856 | 3.072 | 17.963 | 0.879 | 280.453 |
|          | CNMF [14,23] | 28.487 | 10.201 | 4.090 | 7.662 | 0.831 | 53.476 |
|          | GSA [20] | 33.503 | 5.774 | 2.336 | 10.498 | 0.899 | 1.083 |
|          | MAPSMM [22] | 27.218 | 11.368 | 4.439 | 7.836 | 0.869 | 154.387 |
|          | SFIMHS [21] | 26.762 | 12.06 | 4.727 | 6.787 | 0.883 | 0.276 |
|          | Ours | 38.079 | 3.829 | 1.527 | 12.174 | 0.955 | 13.835 |
| SRF2 PSF2 | Hysure [13] | 29.335 | 9.033 | 3.441 | 18.692 | 0.856 | 279.973 |
|          | CNMF [14,23] | 28.049 | 10.708 | 4.306 | 7.914 | 0.835 | 52.585 |
|          | GSA [20] | 32.609 | 6.346 | 2.559 | 12.127 | 0.868 | 1.400 |
|          | MAPSMM [22] | 26.273 | 12.679 | 4.946 | 8.377 | 0.847 | 158.632 |
|          | SFIMHS [21] | 25.75  | 13.566 | 5.312 | 7.485 | 0.864 | 0.391 |
|          | Ours | 36.224 | 4.705 | 1.900 | 12.475 | 0.947 | 13.384 |
Figure 2. Fusion results on the Pavia University dataset (formed by bands 30th, 60th, and 90th). First row, SRF1,PSF1 case; fusion image of the comparison method; Second row, SRF2,PSF1 case, fusion image of the comparison method; Third row, SRF2, PSF2 case, fusion image of the comparison method [13,14,20–23].

Figure 3. Fusion results for oil_painting in the CAVE dataset (formed by bands 10th, 19th, and 28th). First row, SRF1,PSF1 case, fusion image of the comparison method; Second row, SRF2,PSF1 case, fusion image of the comparison method; Third row, SRF2, PSF2 case, fusion image of the comparison method [13,14,20–23].
6.3. Practicality of the Proposed Method

The superiority of blind fusion algorithm fusion performance is finally reflected in the fusion of the real dataset. Since the real dataset does not provide a reference image to calculate the objective fusion metrics of the fused and reference images like the simulated dataset, we demonstrate the practicality of our proposed method by comparing the subjective results of the fused images in this subsection.

Figure 4 shows the fusion results of the comparison methods on the real dataset, and we enlarge the red box to show the top left corner. From the figure, the fused images of MSPSMM and SFIMHS have large blurred patches, and the fused images of Hysure, CNMF and GSA methods contain more streaks, only the fused images of our proposed method are similar to HR-MSI, and the fused images are relatively smooth.

Figure 4. The fusion results of the compared methods on the real dataset [13,14,20–23].

7. Conclusions

This paper first proposes an SRF estimation method based on the quadratic programming problem and multiplicative iterative optimization and uses two advanced semi-blind algorithms to test the accuracy of our estimated SRF. In order to further realize blind fusion, we propose a fusion technology that only needs to provide SRF, using our estimated SRF combined with our fusion algorithm, thus achieving a blind fusion technology that does not rely on SRF and PSF. We have conducted experiments on both the remote sensing dataset Pavia University and the ground dataset CAVE. The subjective and objective results of the experiment show the superior performance of our proposed method. In addition, we also fused the real remote sensing images obtained by the two satellite-borne spectroscopic cameras of sentinel 2 and Hyperion, and the subjective results of the fused images showed the practicality of our proposed method.

Author Contributions: Conceptualization, Y.P. and J.L.; methodology, Y.P. and J.L.; software, J.L.; validation, J.L.; data curation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, Y.P.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the National Natural Science Foundation of China (No. 91948303-1, No. 61803375, No. 12002380, No. 62106278, No. 62101575, No. 61906210) and the National University of Defense Technology Foundation (No. ZK20-52).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Sample Availability: The code in the articles is available from the authors.
References

1. Samiappan, S. Spectral Band Selection for Ensemble Classification of Hyperspectral Images with Applications to Agriculture and Food Safety. Ph.D. Thesis, Gradworks, Mississippi State University: Starkwell, MS, USA, 2014.

2. Liu, Z.; Yan, J.Q.; Zhang, D.; Li, Q.L. Automated tongue segmentation in hyperspectral images for medicine. *Appl. Opt.* **2007**, *46*, 8328–8334. [CrossRef]

3. Pechaneč, V.; Mráz, A.; Rozkošný, L.; Vyvlec̆ka, P. Usage of Airborne Hyperspectral Imaging Data for Identifying Spatial Variability of Soil Nitrogen Content. *ISPRS Int. J. Geo-Inf.* **2021**, *10*. [CrossRef]

4. Meer, F.D.V.D.; Werff, H.M.A.V.D.; Ruitenbeek, F.J.A.V.; Hecker, C.A.; Bakker, W.H.; Nooßen, M.F.; Meijde, M.V.D.; Carranza, E.J.M.; Smith, J.B.D.; Woldai, T. Multi- and hyperspectral geologic remote sensing: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14*, 112–128. [CrossRef]

5. Cui, Y.; Zhang, B.; Wang, W.; Yi, X.; Tang, Y. Deep CNN-based Visual Target Tracking System Relying on Monocular Image Sensing. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018.

6. Cui, Y.; Zhang, B.; Wang, Z.; Li, Y.; Yi, X.; Tang, Y. End-to-End Visual Target Tracking in Multi-Robot Systems Based on Deep Convolutional Neural Network. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 1113–1121.

7. Li, S.; Dian, R.; Fang, L.; Bioucas-Dias, J.M. Fusing Hyperspectral and Multispectral Images via Coupled Sparse Tensor Factorization. *IEEE Trans. Image Process.* **2018**, *27*, 4118–4130. [CrossRef] [PubMed]

8. Dian, R.; Li, S. Hyperspectral Image Super-Resolution via Subspace-Based Low Tensor Multi-Rank Regularization. *IEEE Trans. Image Process.* **2019**, *28*, 5135–5146. [CrossRef] [PubMed]

9. Long, J.; Peng, Y.; Li, J.; Zhang, L.; Xu, Y. Hyperspectral Image Super-resolution via Subspace-based Fast Low Tensor Multi-Rank Regularization. *Infrared Phys. Technol.* **2021**, *103*, 106361. [CrossRef]

10. Long, J.; Peng, Y.; Zhou, T.; Zhao, L.; Li, J. Fast and Stable Hyperspectral Multispectral Image Fusion Technique Using Moore—Penrose Inverse Solver. *Appl. Sci.* **2021**, *11*. [CrossRef]

11. Kanatsoulis, C.I.; Fu, X.; Sidiropanos, N.D.; Ma, W. Hyperspectral Super-Resolution: A Coupled Tensor Factorization Approach. *IEEE Trans. Signal Process.* **2018**, *66*, 6503–6517. [CrossRef]

12. Dian, R.; Li, S.; Fang, L.; Lu, T.; Bioucas-Dias, J.M. Nonlocal Sparse Tensor Factorization for Semiblind Hyperspectral and Multispectral Image Fusion. *IEEE Trans. Cybern.* **2019**, *50*, 4469–4480. [CrossRef]

13. Simoes, M.; Bioucas-Dias, J.; Almeida, L.B.; Chanussot, J. A Convex Formulation for Hyperspectral Image Superresolution via Subspace-Based Regularization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3373–3388. [CrossRef]

14. Yokoya, N.; Mayumi, N.; Iwasaki, A. Cross-Calibration for Data Fusion of EO-1/Hyperion and Terra/ASTER. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 419–426. [CrossRef]

15. Carper, W.J.; Lillesand, T.M.; Kiefer, P.W. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 459–467.

16. Tu, T.M.; Huang, P.S.; Hung, C.L.; Chang, C.P. A Fast Intensity–Hue–Saturation Fusion Technique With Spectral Adjustment for IKONOS Imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 309–312. [CrossRef]

17. Gonzalez-Audicana, M.; Otazu, X.; Fors, O.; Alvarez-Mozos, J. A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1683–1691. [CrossRef]

18. Zhang, Y. Understanding image fusion. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 657–661.

19. Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener. U.S. Patent No. 6,011,875, 4 January 2000.

20. Aiazzi, B.; Baronti, S.; Selva, M. Improving Component Substitution Pansharpening Through Multivariate Regression of MS + Pan Data. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3230–3239. [CrossRef]

21. Smoothing Filter-based Intensity Modulation: A spectral preserve image fusion technique for improving spatial details. *Int. J. Remote Sens.* **2000**, *21*, 3461–3472. [CrossRef]

22. Eismann, M.T. Resolution Enhancement of Hyperspectral Imagery Using Maximum a Posteriori Estimation with a Stochastic Mixing Model. Ph.D. Thesis, University of Dayton, Dayton, OH, USA, 2004.

23. Yokoya, N.; Member, S.; Yairi, T.; Iwasaki, A. Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 528–537. [CrossRef]

24. Frederic, P.M.; Dufaux, F.; Winkler, S.; Ebrahimi, T.; Sa, G. A No-Reference Perceptual Blur Metric. In Proceedings of the International Conference on Image Processing, Rochester, NY, USA, 22–25 September 2002; Volume 3, pp. III–57–III–60.

25. Dell’Acqua, F.; Gamba, P.; Ferrari, A.; Palmason, J.A.; Benediktsson, J.A.; Arnason, K. Exploiting spectral and spatial information in hyperspectral urban data with high resolution. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 322–326. [CrossRef]

26. Yasuma, F.; Mitsunaga, T.; Iso, D.; Nayar, S.K. Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range, and Spectrum. *IEEE Trans. Image Process.* **2010**, *19*, 2241–2253. [CrossRef] [PubMed]

27. Wei, Q.; Bioucas-Dias, J.; Dobigeon, N.; Tourneret, J.Y. Hyperspectral and Multispectral Image Fusion based on a Sparse Representation. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3658–3668. [CrossRef]
28. Wald, L. Quality of high resolution synthesised images: Is there a simple criterion? In *Third Conference Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images; SEE/URISCA: Nice, France, 2000*; pp. 99–103.

29. Yokoya, N.; Grohnfeldt, C.; Chanussot, J. Hyperspectral and multispectral data fusion: A comparative review of the recent literature. *IEEE Geosci. Remote Sens. Mag.* 2017, 5, 29–56. [CrossRef]