



Article

A Spatial-Enhanced LSE-SFIM Algorithm for Hyperspectral and Multispectral Images Fusion

Yulei Wang ^{1,2} , Qingyu Zhu ¹, Yao Shi ¹, Meiping Song ^{1,*} and Chunyan Yu ¹

¹ Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College, Dalian Maritime University, Dalian 116026, China; wangyulei@dmlu.edu.cn (Y.W.); zhuqingyu@dmlu.edu.cn (Q.Z.); 1120180233@dmlu.edu.cn (Y.S.); yucy@dmlu.edu.cn (C.Y.)

² State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710000, China

* Correspondence: smping@dmlu.edu.cn

Abstract: The fusion of a hyperspectral image (HSI) and multispectral image (MSI) can significantly improve the ability of ground target recognition and identification. The quality of spatial information and the fidelity of spectral information are normally contradictory. However, these two properties are non-negligible indicators for multi-source remote-sensing images fusion. The smoothing filter-based intensity modulation (SFIM) method is a simple yet effective model for image fusion, which can improve the spatial texture details of the image well, and maintain the spectral characteristics of the image significantly. However, traditional SFIM has a poor effect for edge information sharpening, leading to a bad overall fusion result. In order to obtain better spatial information, a spatial filter-based improved LSE-SFIM algorithm is proposed in this paper. Firstly, the least square estimation (LSE) algorithm is combined with SFIM, which can effectively improve the spatial information quality of the fused image. At the same time, in order to better maintain the spatial information, four spatial filters (mean, median, nearest and bilinear) are used for the simulated MSI image to extract fine spatial information. Six quality indexes are used to compare the performance of different algorithms, and the experimental results demonstrate that the LSE-SFIM based on bilinear (LES-SFIM-B) performs significantly better than the traditional SFIM algorithm and other spatially enhanced LSE-SFIM algorithms proposed in this paper. Furthermore, LSE-SFIM-B could also obtain similar performance compared with three state-of-the-art HSI-MSI fusion algorithms (CNMF, HySure, and FUSE), while the computing time is much shorter.

Keywords: hyperspectral image; multi-source image fusion; SFIM; least square estimation; spatial filter



Citation: Wang, Y.; Zhu, Q.; Shi, Y.; Song, M.; Yu, C. A Spatial-Enhanced LSE-SFIM Algorithm for Hyperspectral and Multispectral Images Fusion. *Remote Sens.* **2021**, *13*, 4967. <https://doi.org/10.3390/rs13244967>

Academic Editors: Chein-I Chang, Haoyang Yu, Jiaojiao Yu, Lin Wang, Hsiao-Ch Li and Xiaorun Li

Received: 14 October 2021

Accepted: 3 December 2021

Published: 7 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, a large number of remote-sensing satellites have been launched continuously with the development of Earth observation technology [1,2]. Modern remote-sensing technology has reached a new developmental stage of multi-platform, multi-sensor, and multi-angle observation [3–5]. The continuous development of remote-sensing applications such as geological exploration [6], resource and environmental investigation [7–9], agricultural monitoring [10–12], urban planning [13–16], etc., has greatly promoted the demand for remote-sensing data and the improvement of the performance of satellite sensors. However, due to the limitations of optical diffraction, modulation transfer function, signal-to-noise ratio, and the sensor hardware conditions, a single sensor normally cannot obtain data with both high-spatial and high-spectral resolutions at the same time. Multi-sensor data fusion has arisen at an historic moment, which can effectively explore the complementary information from multi-platform observations, making land surface monitoring more accurate and comprehensive. Multi-source remote-sensing data fusion refers to the processing of multi-source data with complementary information in time or space according to certain rules, so as to obtain a more accurate and informative composite images than any single data source.

A variety of multi-source remote-sensing fusion techniques have been developed in the last decade to enhance the spatial resolution of hyperspectral images and obtain information-rich HSI data with both high-spectral and high-spatial resolutions. The HSI-MSI fusion algorithms can be divided into the following four categories: component substitution (CS), multi-resolution analysis (MRA), spectral unmixing (SU), and Bayesian representation (BR). The idea of CS-based fusion algorithms is straightforward, it transforms the original HSI data, replaces the spatial information in low-spatial HSI data set with the spatial information in the high-spatial MSI data set, and finally inverts the reconstructed data to obtain the fused hyperspectral image. The typical CS-based methods are proposed toward generalizing existing pansharpening methods for HSI-MSI fusion, including the IHS [17] transform method proposed by W. J. Carper in 1990, PCA [18] transform proposed by P. S. Chavez in 1991, Gram-Schmidt (GS) [19] transform proposed by B. Aiazzi in 2007, and their variants [20–22], etc. These methods are simple and easy to implement, but have serious spectral distortion and cannot be used well for the fusion of hyperspectral images. The MRA-based methods obtain the fusion result by filtering the high-resolution image and adding the high-frequency detailed information to the hyperspectral image. The earliest MRA-based methods realized multi-scale image decomposition through pyramid transform, while the most representative and most widely used multi-resolution analysis methods include the fusion method based on various wavelet transforms [23], the smoothing filter-based intensity modulation (SFIM) [24] proposed by Liu and the generalized Laplacian pyramid (GLP) method proposed by Aiazzi [25]. The advantages of these MRA-based methods are less spectral distortion and anti-aliasing, but the algorithm is complex and the spatial feature loss often occurs in the fusion results. The SU-based methods utilize the hyperspectral linear unmixing model and apply it to the fusion optimization model. The advantages of these methods are less spatial and spectral information loss, but they always have higher computational complexity. Typical methods include the coupled non-negative matrix factorization (CNMF) [26] method proposed by Yokoya in 2012, the subspace-based regularization (HySure) [27] method proposed by Simoes in 2015, and the coupled sparse tensor factorization (CSTF) [28] method proposed by Li in 2018. The BR-based methods transform the problem of high-resolution image and hyperspectral image fusion into the problem of solving the Bayesian optimization model, and obtain the fusion result through solving the optimization. Typical BR-based methods include a maximum posteriori-stochastic mixing model (MAP-SMM) [29] proposed by Eismann in 2004, Bayesian sparse method [30] proposed by Wei in 2015, and fast fusion based on the Sylvester equation (FUSE) [31] method proposed by Wei in 2015. The BR-based methods have the advantage of less spatial and spectral information loss, but also result in a disadvantage of high computational complexity.

Recently, an increasing number of HSI-MSI fusion algorithms have been proposed [32–34]. These algorithms have been proved to be effective with good fusion performance. However, most of the researchers focus too much on performance improvements using modern technologies such as sparse representation, deep learning processing, etc., ignoring the computing time. In other words, these algorithms improve the fusion performance at the cost of increased computational complexity. As one of the effective MRA-based fusion methods, SFIM is proposed by Liu [24] for image fusion as mentioned above. Compared with traditional methods, SFIM is simple to calculate, easy to implement, and the spectral information is normally retained well, but there are problems such as fuzzy edge information of the image and insufficient improvement of detailed spatial information. In recent years, many improved SFIM algorithms have been studied, most of which focused on how to obtain simulated multispectral images with spatial information characteristics consistent with hyperspectral images and spectral features consistent with multispectral images. This paper combines the least square estimation LSE algorithm with SFIM, which can effectively improve the spatial information quality of the fused image. This paper also compares several spatial filters to extract spatial information to enhance the simulated MSI's boundary spatial information, and proposes an improved LSE-SFIM fusion algorithm based on

spatial information promotion to obtain an optimal fusion result. Experimental results on three HSI-MSI data sets show the effectiveness of the proposed algorithm using six image quality indexes.

The remainder of this article is organized as follows. Section 2 gives a detailed description of the proposed method. In Section 3, experimental results and analysis of different data sets are presented. Finally, conclusions are drawn in Section 4.

2. Proposed Method

2.1. Basic Smoothing Filter-Based Intensity Modulation (SFIM) Algorithm

The SFIM algorithm was proposed by Liu for image fusion in 2000, which is based on the simplified solar radiation and surface reflection model. Even though it was proposed some time ago, this algorithm is still in use due to its simplicity with good spectral preservation. The basic principle of the traditional SFIM is expressed as follows [24]:

$$DN_{SFIM} = \frac{DN_{low}DN_{high}}{MeanDN_{high}} \quad (1)$$

where DN_{low} , DN_{high} represent the gray values of low-resolution and high-resolution images respectively, $MeanDN_{high}$ represents the simulated low-resolution image obtained by the local mean value of DN_{high} .

2.2. The Proposed Spatial Filter-Based Least Square Estimation (LSE)-SFIM

For HSI-MSI image fusion, the formula (1) can be expressed as:

$$Fusion = \frac{HSI' \times MSI}{MSI''} \quad (2)$$

where HSI' is the up-sampling of original low-resolution hyperspectral data HSI, MSI represents the original high-resolution multispectral data MSI, and MSI'' is the up-sampling of MSI' where MSI' represents the simulated low-resolution image obtained by MSI. The algorithm performance is influenced by two factors: (1) how to obtain the simulated low-resolution image MSI' ; and (2) how to get the up-sampled HSI' and MSI'' . The traditional SFIM uses mean filter to obtain the simulated low-resolution MSI' (down-sampling) and uses the same filter to obtain up-sampled HSI' and MSI'' . The edge information is lost by the mean filters, and this paper takes two steps to solve the problem: (1) least squares estimation (LSE) is used to adjust the coefficient and obtain MSI' with as similar spatial information as the original HSI image, with the details discussed in Section 2.2.1; (2) filtering and interpolation methods are compared in the up-sampling stage to obtain the best enhanced spatial information, with the details discussed in Section 2.2.2. A bilinear approach proved to be the best in the experiments for this paper. The flow chart of the proposed algorithm is shown in Figure 1.

In order to make it clear how we obtain MSI , MSI' , MSI'' , HSI , and HSI' , Figure 2 gives a graphic abstract with detailed steps of the proposed algorithm. It can easily be seen from Figure 2 that, MSI is the original high-spatial multispectral data set, MSI' is the down-sampling of MSI where the spatial size can be shrunk into the same as the original HSI (LSE is used here to adjust MSI' for preserving better spatial information), and MSI'' is the up-sampling of MSI' with the same size as the original high-spatial resolution MSI . HSI is the original low-spatial resolution hyperspectral data set, and HSI' is the up-sampling of HSI to the same size as the high-spatial resolution.

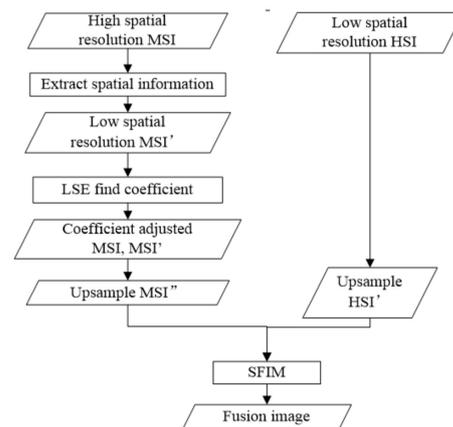


Figure 1. Flowchart of the proposed fusion algorithm.

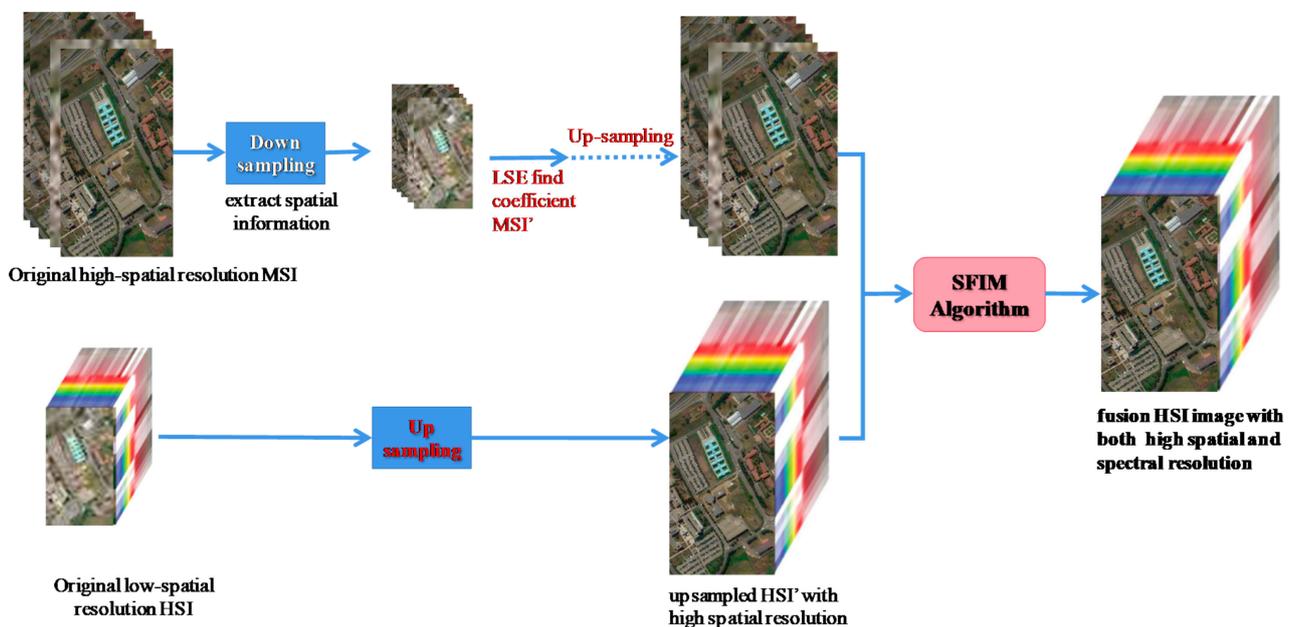


Figure 2. Graphic abstract with detailed steps of the proposed algorithm.

2.2.1. Least Square Estimation Based SFIM Algorithm (LSE-SFIM)

Assuming that there is an ideal simulated multispectral image MSI' , it should have two characteristics: (1) the spatial information characteristics are consistent with the original hyperspectral image, which can ensure that the spatial information of the hyperspectral image is counteracted; (2) the spectral information characteristics should be consistent with the original multispectral image, which ensures that the spectral characteristics of the multispectral image can be counteracted. The least square estimation algorithm solves these two problems well.

The LSE algorithm finds the best matching function by minimizing the sum of squares of errors. It is often used to solve linear regression coefficients in the processing of remote-sensing images. The LSE-SFIM algorithm uses LSE to solve the linear regression coefficient that can minimize the spatial information error between the hyperspectral image and the simulated multispectral image MSI' , so that the latter can have as much spatial information as possible as with the hyperspectral image.

The LSE-SFIM fusion algorithm first down-samples and extracts the spatial information of the high-resolution multispectral image MSI to obtain the simulated MSI' , and then uses the least square estimation algorithm to solve the problem that can minimize the linear

regression coefficient of the spatial information error between the multispectral image MSI' and the hyperspectral image, and use this linear regression coefficient to update the MSI and MSI', so as to ensure that the spectral information of the MSI is close to the MSI', and it can ensure that both MSI and MSI' can have the same spatial information as HSI as possible. Finally, the HSI' and MSI'' are obtained by up-sampling, which will be introduced in the next section, and fused separately according to the bands to obtain the fused image.

2.2.2. Spatial Information Enhanced LSE-SFIM

When using the LSE-SFIM fusion image, the most critical step concerns how to obtain the simulated multispectral image MSI'' whose spatial information features are consistent with the hyperspectral image and spectral features are consistent with the multispectral image, so as to effectively improve the spatial resolution of the hyperspectral image and achieve the purpose of fusion. This step is the up-sampling process to obtain MSI'' and HSI' in Figure 2. This paper compares several methods of extracting boundary spatial information from low-spatial resolution multispectral images, and obtains the best fusion result.

Filtering Method

Mean filtering and median filtering are two commonly used filtering methods. The main idea of mean filtering is to replace the gray value of the central pixel with the mean value of the gray value of the pixel to be found and the surrounding pixels in the middle, so as to achieve the purpose of filtering. Mean filtering can be simplified as Equation (3):

$$g(x, y) = \frac{1}{M} \sum_{f \in W} f(x, y) \quad (3)$$

where M is the filtering window size (pixels within the current window), and W is the current window.

Median filtering, as the name implies, is to replace the value of the pixel with the median value of the gray-scale values in the neighborhood window of a pixel. Median filtering can be simplified as Equation (4):

$$g(x, y) = \text{med}\{f(x - k, y - l), (k, l \in W)\} \quad (4)$$

Taking 3×3 window size as the example, mean and median filtering are shown in Figure 3, where (a) represents gray values before filtering, (b) represents gray values after mean filtering, and (c) represents gray values after median filtering.

1	2	1	4	3
1	2	2	4	3
5	7	6	8	5
4	5	4	7	6
5	9	9	8	1

(a)

1	2	1	4	3
1	3	4	4	3
5	4	5	5	5
4	6	7	6	6
5	9	9	8	1

(b)

1	2	1	4	3
1	2	4	4	3
5	4	5	5	5
4	5	7	6	6
5	9	9	8	1

(c)

Figure 3. 3×3 mean and median filtering, (a) gray values before filtering in green color, (b) gray values after mean filtering in red color, and (c) gray values after median filtering in purple color.

Interpolation Method

Image interpolation algorithm is a basic technology in image processing. Nearest neighbor interpolation and bilinear interpolation are two commonly used image interpolation algorithms. The nearest neighbor interpolation algorithm has the least amount of calculation and the simplest principle. The gray value of the nearest point among the neighboring pixels around the point to be sampled is used as the gray value of the point. There

is a linear relationship between the pixel values of different points in the image. According to this idea, the bilinear interpolation algorithm considers the pixel values in the horizontal and vertical directions at the same time, so that the problem of grayscale discontinuity in the image is improved, and the overall effect of the image can also be improved.

3. Experimental Results and Analysis

In order to verify the effectiveness of the algorithm proposed in this paper, three sets of simulation experiment data sources are selected, namely Pavia University, Chikusei and HyMap Rodalquilar. This paper uses MATLAB R2018b software platform to program on Windows 10 64-bit system, and the processor is Intel (R) Core (TM) i5-8250U, 8G memory.

3.1. Hyperspectral Datasets

In order to evaluate the performance of the fusion method objectively and quantitatively, we use low-spatial resolution hyperspectral images obtained from real data resampling in the spatial domains, and high-spatial resolution multispectral images obtained in the spectral domains to carry out simulation data experiments. Table 1 shows the parameters of the three datasets used in this paper for verification experiments.

Table 1. Parameters of three hyperspectral datasets.

Dataset	Year	Original Sensor	Spectral Range (μm)	Spatial Resolution (m)	Bands
Pavia University	2003	ROSIS-3	0.43–0.84	1.3	103
Chikusei	2014	Hyperspec	0.36–1.02	2.5	128
HyMap Rodalquilar	2003	HyMap	0.4–2.5	10	167

3.1.1. Pavia University

Pavia University acquired a ROSIS sensor in 2001. The image size is 610×340 with a spatial resolution of 1.3 m per pixel, and the experimental data selected in this section contains 560×320 pixels. Its spectral range is 0.43–0.86 μm with a total of 115 bands, and 103 bands are remaining used after removing 12 noise bands. The low-spatial resolution HSI image was obtained from the original HSI data through the isotropic Gaussian point spread function down-sampling eight times with a total of 103 bands and 70×40 pixels, the pseudo-color image is shown in Figure 4a. The high-spatial resolution MSI image data was synthesized from the original HSI data according to the SRF down-sampling of the ROSIS sensor. There were four bands in total with an image size of 560×320 , as shown in Figure 4b. The reference image of the original HIS is shown in Figure 4c.

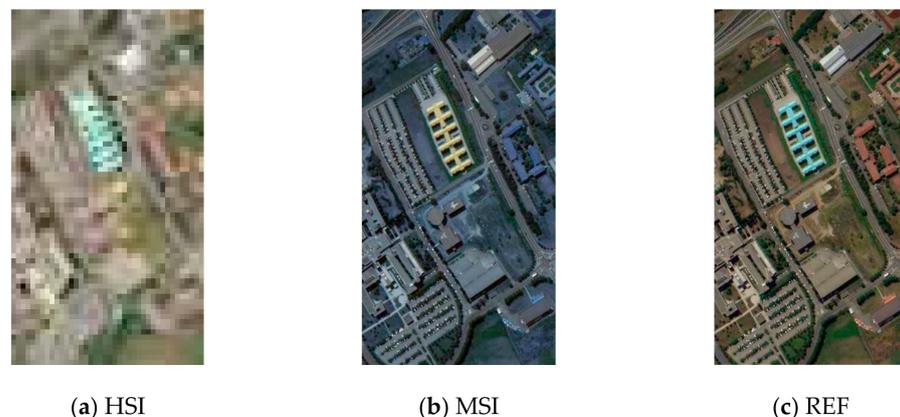


Figure 4. Pavia University datasets with (a) low spatial resolution hyperspectral image (HSI), (b) high spatial resolution multispectral image (MSI) and (c) high spatial resolution HSI as the reference.

3.1.2. Chikusei

Chikusei dataset was collected by a Headwall Hyperspec-VNIR-C sensor on 29 July 2014 in Chikusei City, Japan. It was then produced and published by Naoto Yokoya and Akira Iwasaki of the University of Tokyo [26]. The spatial resolution is 2.5 m and the scene is 2517×2335 . It consists of several pixels, mainly including agricultural and urban areas. In the experiment, a size of 540×420 pixels image is selected for experiments. The data spectrum range is $0.36\text{--}1.02 \mu\text{m}$, including 128 bands in total. Among them, the low-spatial resolution HSI image was obtained from the original HSI data through the isotropic Gaussian point spread function down-sampling six times, with a total of 128 bands and 90×70 pixels, the pseudo-color image is shown in Figure 5a. The high-spatial resolution MSI data were synthesized from the original HSI data according to the SRF of the WV-2 sensor, with eight bands and 540×420 pixels, and the pseudo-color image of MSI was shown in Figure 5b. The reference image of the original HSI is shown in Figure 5c.

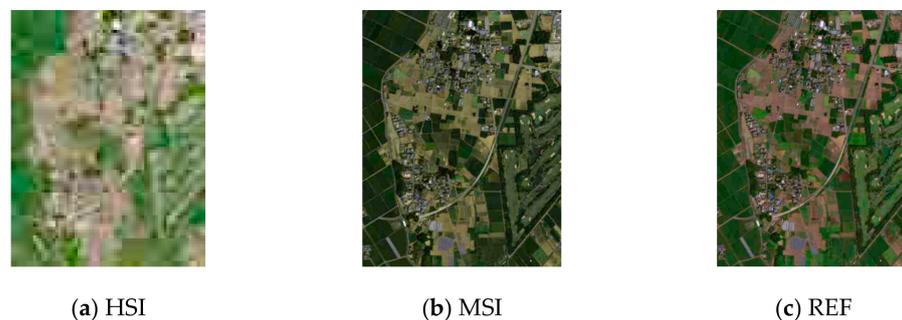


Figure 5. Chikusei datasets with (a) low-spatial resolution HSI, (b) high-spatial resolution MSI and (c) high-spatial resolution HSI as the reference.

3.1.3. HyMap Rodalquilar

The HyMap image was taken in Rodalquilar, Spain in June 2003 [35], covering a gold mining area in the Cabo de Gata Mountains. The spatial resolution of the data was 10 m. The experimental data selected in this paper contains 867×261 pixels. After removing the water absorption band, 167 bands are selected for experimentation, and the spectral range is $0.4 \mu\text{m}\text{--}2.5 \mu\text{m}$. Among them, the low-spatial resolution HSI image was obtained from the original HSI data through the isotropic Gaussian point spread function down-sampling three times, with a total of 167 bands and 289×87 pixels, and the resulting pseudo-color image is shown in Figure 6a. The high-spatial resolution MSI image data were synthesized from the original HSI data according to the SRF down-sampling of the HyMap sensor. There were four bands in total with the image size of 867×261 pixels, as shown in Figure 6b. The reference image is shown in Figure 6c.

3.2. Comparative Analysis of the Proposed Spatial Enhanced LSE-SFIM Using Different Spatial Filters

In this section, different spatial enhanced methods in Section 2.2.2 are used to extract boundary information and obtain better fusion results, and the performance discussed to find the best method. Six methods are discussed in this section, the traditional SFIM (named as SFIM), LSE-based SFIM (named LSE-SFIM), mean filtering LSE-SFIM (named LSE-SFIM-M), median filtering LSE-SFIM (named LSE-SFIM-Med), neighboring interpolation LSE-SFIM (named LSE-SFIM-N), and bilinear interpolation LSE-SFIM (named LSE-SFIM-B). Both subjective and objective evaluations are discussed, and spectral distortion are compared among all six algorithms.

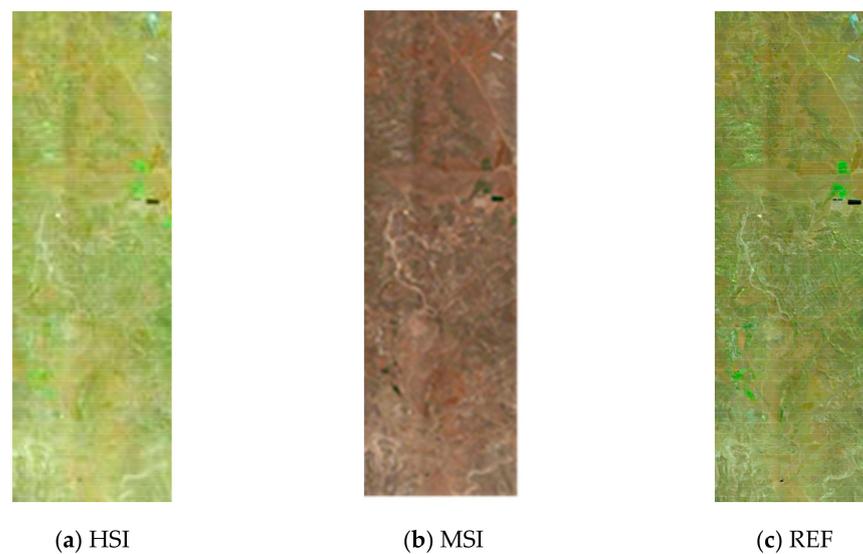


Figure 6. HyMap Rodalquilar datasets with (a) low spatial resolution HSI, (b) high spatial resolution MSI and (c) high spatial resolution HSI as the reference.

3.2.1. Subjective Evaluation

The subjective evaluation mainly uses human eyes to observe the fusion results. The comparison chart of the fusion results of the three groups of experiments is shown in Figures 7–9. Observing the fusion result from a subjective point of view, it can be known that the method of using bilinear interpolation to obtain simulated MSI' has better visibility, and the fusion result obtained has a clearer texture and better spectrum retention performance.

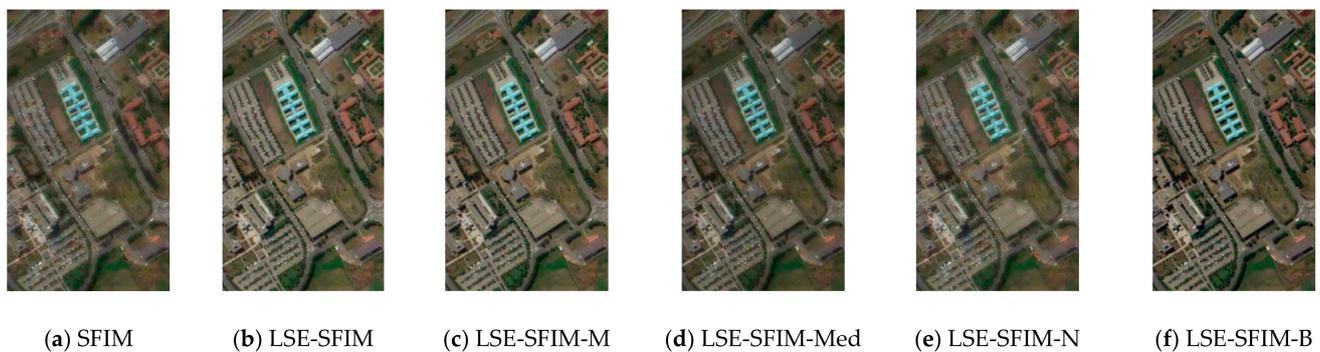


Figure 7. Fusion results of Pavia University using six smoothing filter-based intensity modulation (SFIM)-based algorithms.



Figure 8. Fusion results of Chikusei using six SFIM-based algorithms.

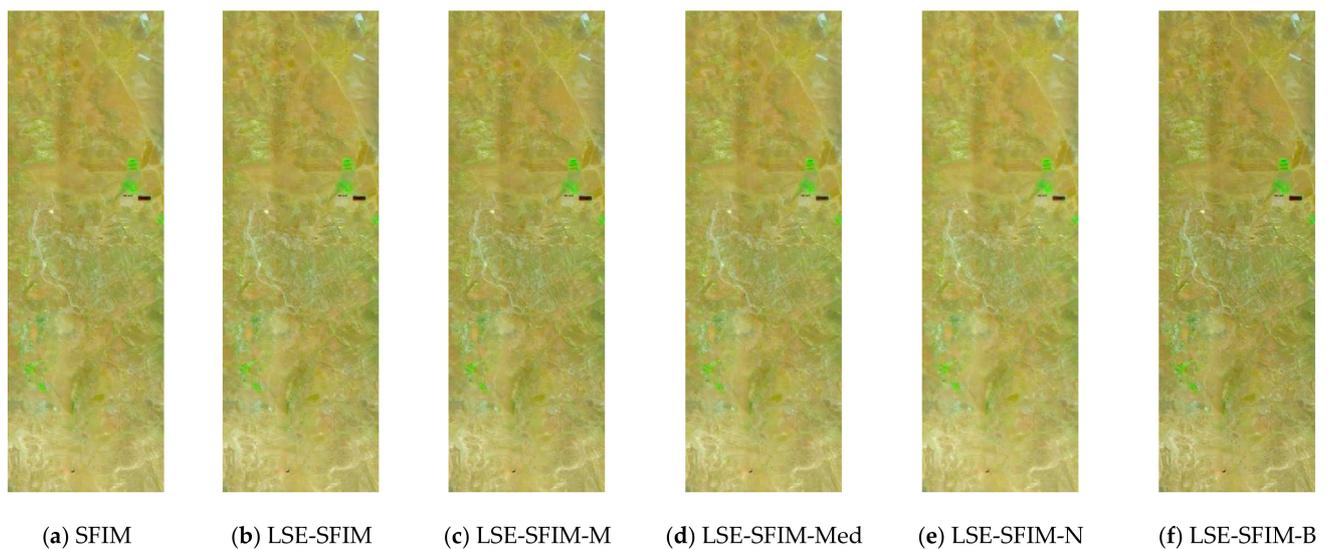


Figure 9. Fusion results of Chikusei using six SFIM-based algorithms.

Figure 7a–f shows that the Pavia University dataset has been subjected to SFIM, LSE-SFIM, LSE-SFIM-M, LSE-SFIM-Med, LSE-SFIM-N and LSE-SFIM-B. It can be seen from the figures that the spectral characteristics of the fusion results by all methods are maintained well. In terms of spatial geometric features, the fusion result of LSE-SFIM-N is visually blurred, and the edge details are not highlighted. In addition, the fusion results obtained by other algorithms have higher clarity, and the LSE-SFIM-B algorithm, whether in terms of spectral characteristics or spatial characteristics, has the closest visual effect to the reference image.

Figure 8a–f are the fusion results of Chikusei dataset through 6 algorithms, namely SFIM, LSE-SFIM, LSE-SFIM-M, LSE-SFIM-Med, LSE-SFIM-N and LSE-SFIM-B. It is obvious that, in terms of spectral characteristics, the fusion results of each method have no more spectral distortions of ground objects, and the color information performs well. In terms of spatial features, the fusion result of LSE-SFIM-N has unclear ground textures and non-obvious edge details. In addition, the fusion results obtained by other algorithms maintain both texture features and edge details of the ground features, especially the LSE-SFIM-B algorithm, which retains more spatial features and the edges of the ground features are more obvious.

Figure 9a–f are the fusion results of the HyMap Rodalquilar data set by 6 algorithms, namely SFIM, LSE-SFIM, LSE-SFIM-M, LSE-SFIM-Med, LSE-SFIM-N and LSE-SFIM-B. In terms of spectral characteristics, the fusion results of each method are not too distorted in the maintenance of the ground object spectrum, and the color information is maintained well. In terms of spatial features, the LSE-SFIM-N fusion result has a poor spatial information enhancement effect. In addition, the fusion results obtained by other algorithms maintain the texture features of hyperspectral images and multispectral images well, especially the LSE-SFIM-B algorithm, which maintains the spatial characteristics better, and the edges of the features are also the most obvious.

In general, through the subjective evaluation of the fusion results by human eyes, it can be found that the proposed LSE-SFIM-B fusion algorithm has the best performance, and the fusion image with clearer boundaries can be obtained from the visual results, especially in Figures 7 and 8. It is the LSE-SFIM-B algorithm makes full use of the complementary characteristics of HIS and MSI images, which realizes the fusion of spectral and spatial features of multiple source images, improves the geometric features of ground objects, and verifies the effectiveness of this algorithm. As for Figure 9, due to image abbreviation, the spatial information enhancement effect of some images is not easy to see, and it is difficult to subjectively judge which method is better. Therefore, objective evaluation is particularly important.

3.2.2. Objective Evaluation

By observing the fusion results in Figures 7–9, it can be seen that the method of obtaining simulated multispectral images by using the LSE-SFIM-B method has better visibility, and the fusion results obtained have clearer texture and better spectrum retention capabilities. To further objectively evaluate the quality of the fusion images by different algorithms, this paper also calculates and analyzes the fusion results from a quantitative perspective by comparing the six objective evaluation indicators of PSNR (peak signal-to-noise ratio), SAM (spectral angle mapping), CC (cross correlation), $Q2^n$ (quality 2ⁿ), RMSE (root mean square error) and ERGAS (error relative global adimensionnelle synthesizer). The quantitative comparisons are shown in Tables 2–4, with the histogram comparison of evaluation indicators shown in Figures 10–12.

Table 2. Quantitative comparisons of fusion performance by six algorithms for Pavia University.

Evaluation Index	SFIM	LSE-SFIM	LSE-SFIM-M	LSE-SFIM-Med	LSE-SFIM-N	LSE-SFIM-B
PSNR	25.8772	36.1762	37.9836	33.4864	28.7486	42.1976
SAM	9.3271	3.4442	3.0933	4.1955	5.7274	2.6762
CC	0.82418	0.98594	0.99101	0.97659	0.92191	0.99362
$Q2^n$	0.46532	0.72695	0.75624	0.7093	0.56614	0.8975
RMSE	0.4142	0.01298	0.01075	0.01767	0.030804	0.007333
ERGAS	6.1140	1.2945	1.0702	1.6944	2.8878	0.76253

Table 3. Quantitative comparison of fusion results by six algorithms for Chikusei.

Evaluation Index	SFIM	LSE-SFIM	LSE-SFIM-M	LSE-SFIM-Med	LSE-SFIM-N	LSE-SFIM-B
PSNR	24.0379	37.8579	40.1123	35.3821	31.0431	46.6653
SAM	7.4477	1.8777	1.5704	2.0696	2.8164	1.3432
CC	0.76329	0.9873	0.99117	0.98013	0.94795	0.99341
$Q2^n$	0.35951	0.87498	0.87572	0.85253	0.83137	0.91992
RMSE	0.4142	0.0078549	0.0061514	0.010291	0.017394	0.0037586
ERGAS	6.1140	1.7005	1.4777	2.0937	3.0949	1.2483

Table 4. Quantitative comparison of fusion results by six algorithms for HyMap Rodalquilar.

Evaluation Index	SFIM	LSE-SFIM	LSE-SFIM-M	LSE-SFIM-Med	LSE-SFIM-N	LSE-SFIM-B
PSNR	36.9969	36.3762	37.8249	36.518	35.2463	39.6276
SAM	2.9165	2.7045	2.6616	2.6922	2.7101	2.6475
CC	0.95908	0.96912	0.97943	0.97128	0.96045	0.9855
$Q2^n$	0.67894	0.51703	0.54971	0.53345	0.49201	0.6217
RMSE	0.018907	0.016785	0.015491	0.01656	0.018003	0.014377
ERGAS	4.2584	2.3641	2.1552	2.3246	2.5765	1.9779

According to Table 2 and Figure 10, it can be seen that for Pavia University data, the method based on the LSE-SFIM algorithm has better fusion effect than traditional SFIM, and the algorithms LSE-SFIM-M and LSE-SFIM-B are better than the original LSE-SFIM algorithm, and the effect of LSE-SFIM-B is even better than that of LSE-SFIM-M, which is the best among several comparison methods. In terms of PSNR, CC and $Q2^n$, the LSE-SFIM-B fusion algorithm is significantly higher than the results of the other algorithms, indicating that the spatial quality information of the fusion image is better, the fusion result has more detailed spatial information, and it is correlated with the reference image. In terms of SAM, RMSE and ERGAS, the LSE-SFIM-B fusion algorithm is still superior to other algorithms, indicating that the fusion result can better maintain the spectrum, and the error with the reference image is the smallest, and the fusion result is the closest to the reference image.

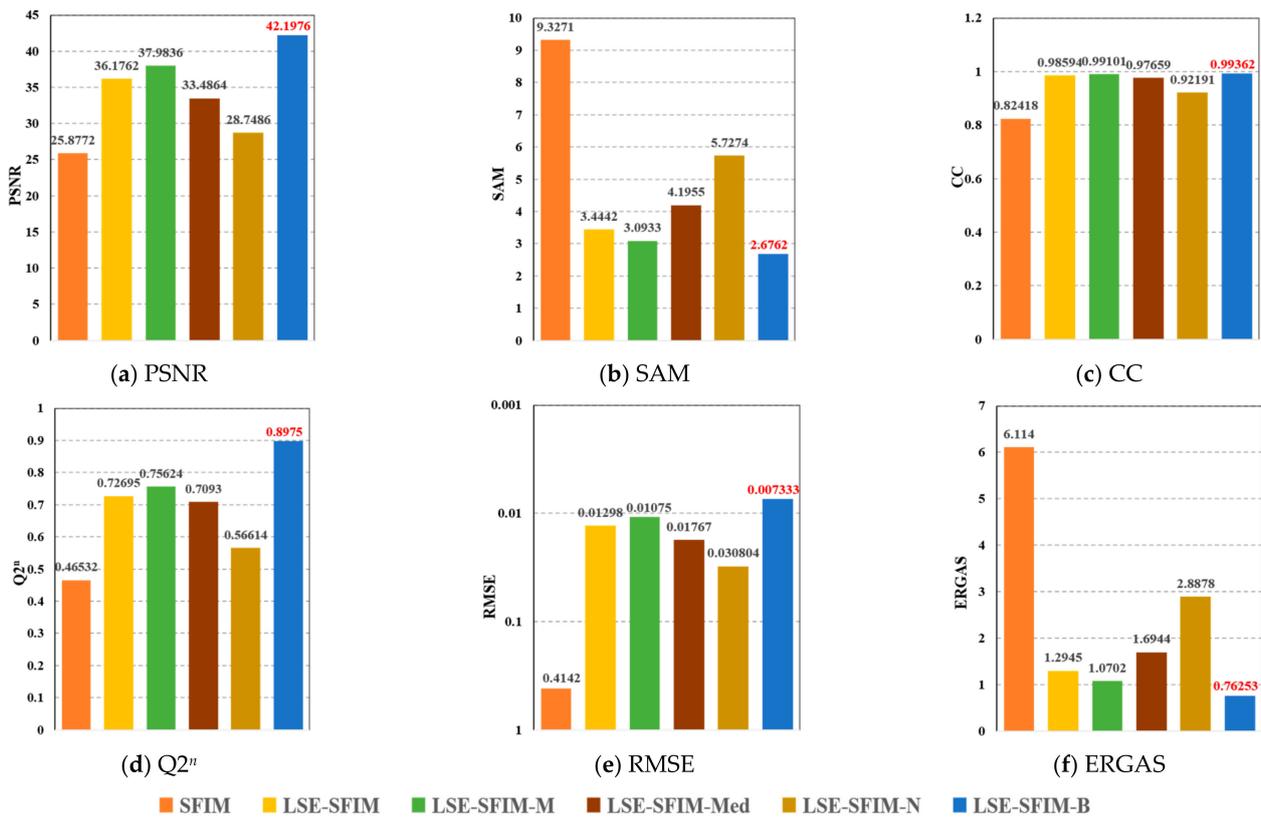


Figure 10. Histograms comparison of evaluation indicators by six algorithms for Pavia University.

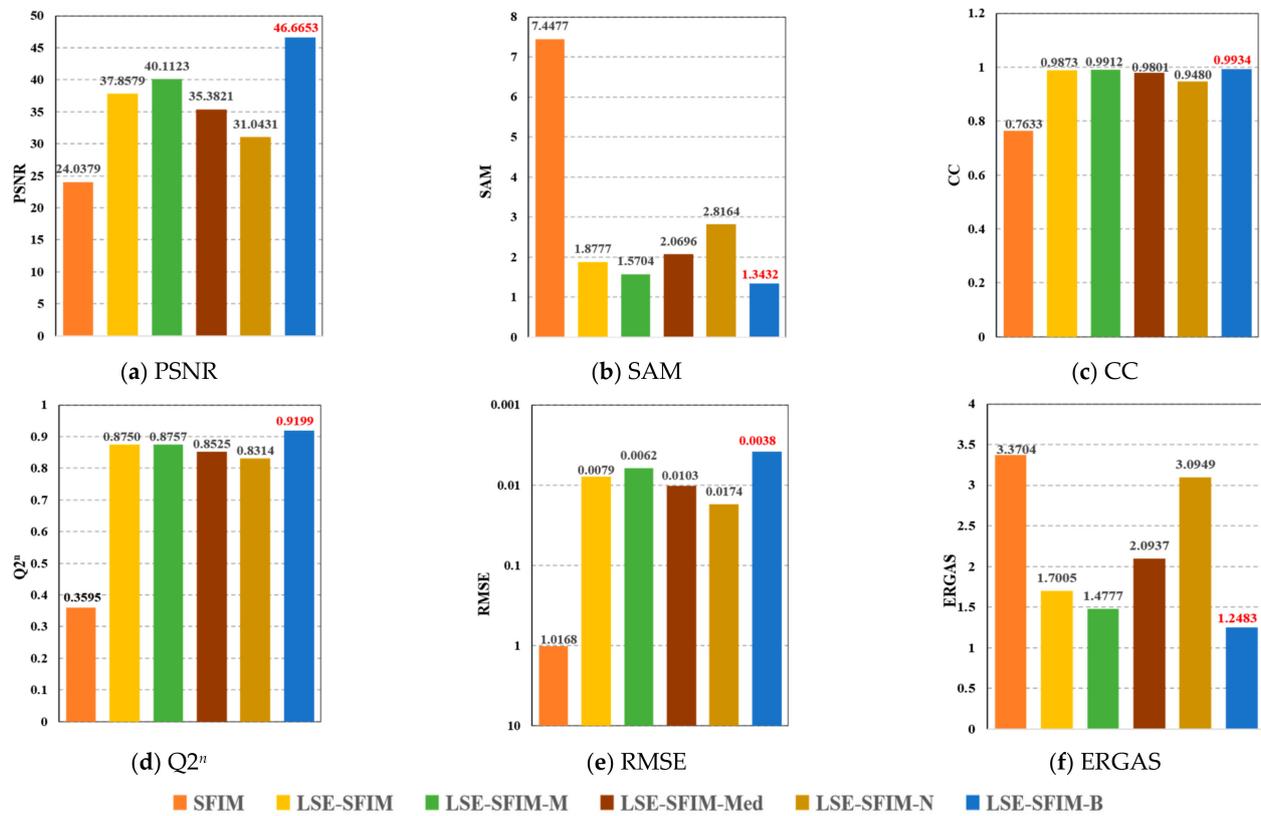


Figure 11. Histogram comparison of evaluation indicators by six algorithms for Chikusei.

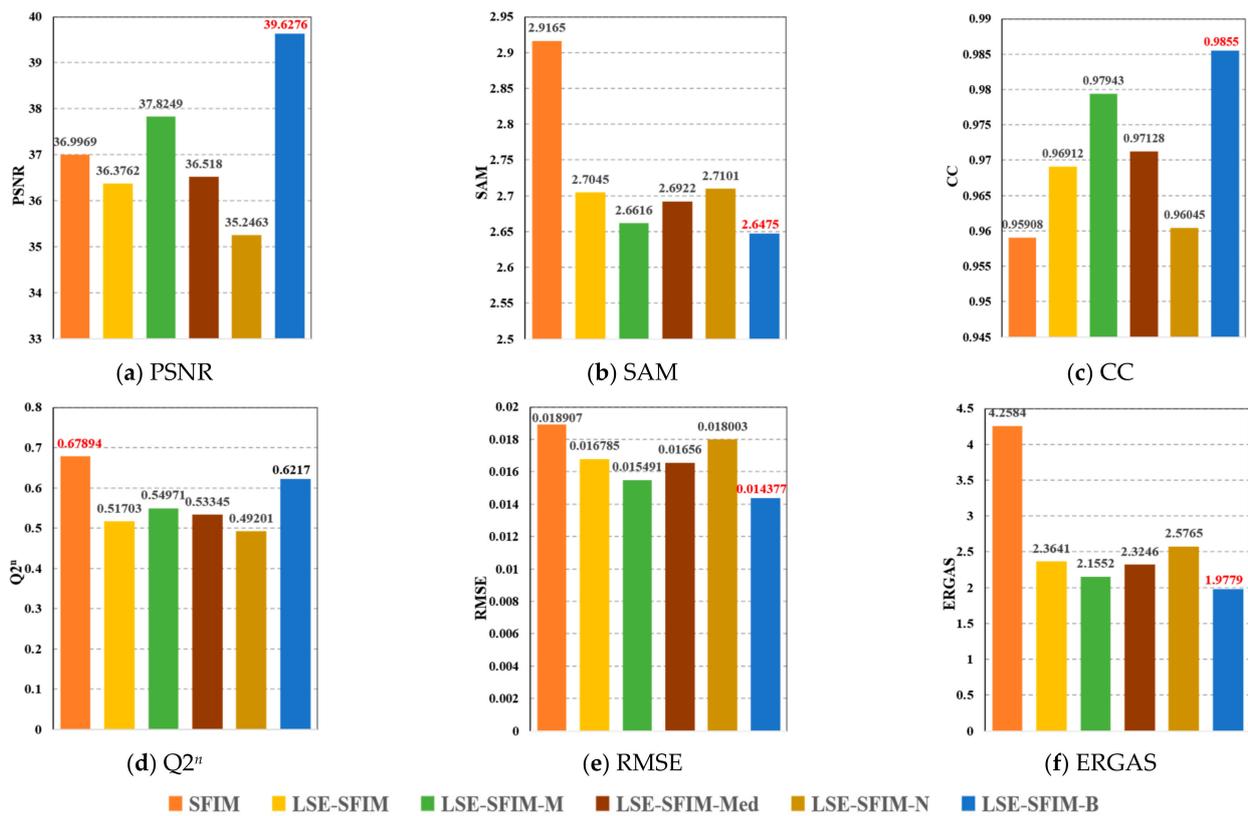


Figure 12. Histogram comparison of evaluation indicators by six algorithms for HyMap Rodalquilar.

In order to further illustrate that the algorithms proposed in this paper have a good performance using the data of different images, Table 3 and Figure 11 give the objective evaluation of Chikusei data, and Table 4 and Figure 12 give the objective evaluation of HyMap Rodalquilar data. It can be seen that for all the data sets, the fusion algorithm using LSE-SFIM-B is also the most outstanding in terms of spatial resolution enhancement and spectral characteristic maintenance. Specifically, the LSE-SFIM-B algorithm is significantly higher than other algorithms in terms of PSNR, CC and $Q2^n$, indicating that the fusion image has good ground quality information, the fusion result is more detailed, and the correlation with the reference image is relatively high. In terms of SAM, RMSE and ERGAS, the LSE-SFIM-B fusion algorithm has the best performance, indicating that the error between the fusion result and the reference image is the smallest, and the spectrum can be better maintained.

3.2.3. Spectral Distortion Comparison

A good fusion method should minimize spectral distortion as much as possible while improving the spatial resolution. In this section, to further analyze the spectral distortion for different SFIM-based algorithms, Figures 13–15 show SAM plots of the experimental results of three hyperspectral data sets. The SAM plot is conducted for every pixel to compute the SAM value between the fusion result and the reference image. In the figures, each pixel uses the change from cold to warm to indicate the level of spectral similarity at that pixel. The closer the color of the pixel point is to the warm color, that is, the closer to dark red, the lower the spectral similarity and the worse the spectral quality relative to other pixels; the closer the color of the pixel point is to the cool color, that is, the closer to dark blue, the higher the spectral similarity and the higher the spectral quality relative to other pixels. The larger the area occupied by the blue part in the figure, the better the overall spectral quality. Compared with other algorithms, it can be seen from the SAM graph of the algorithm experiment results that the spectral performance of LSE-SFIM-B on the three data sets is relatively better.

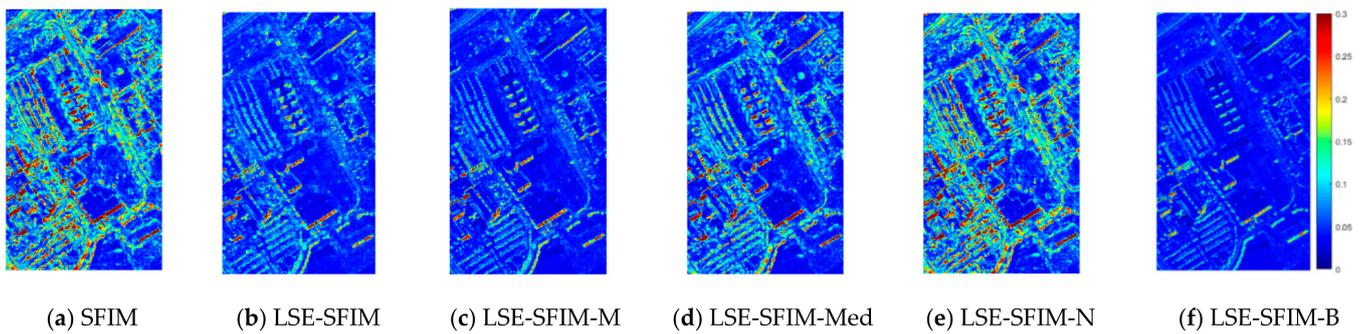


Figure 13. Spectral angle mapping (SAM) map of six algorithms for Pavia University data experiment results.

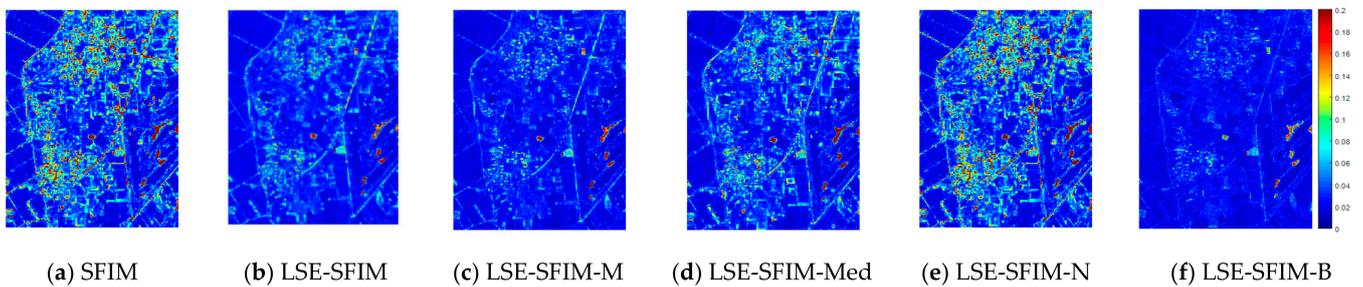


Figure 14. SAM map of six algorithms for Chikusei data experiment results.

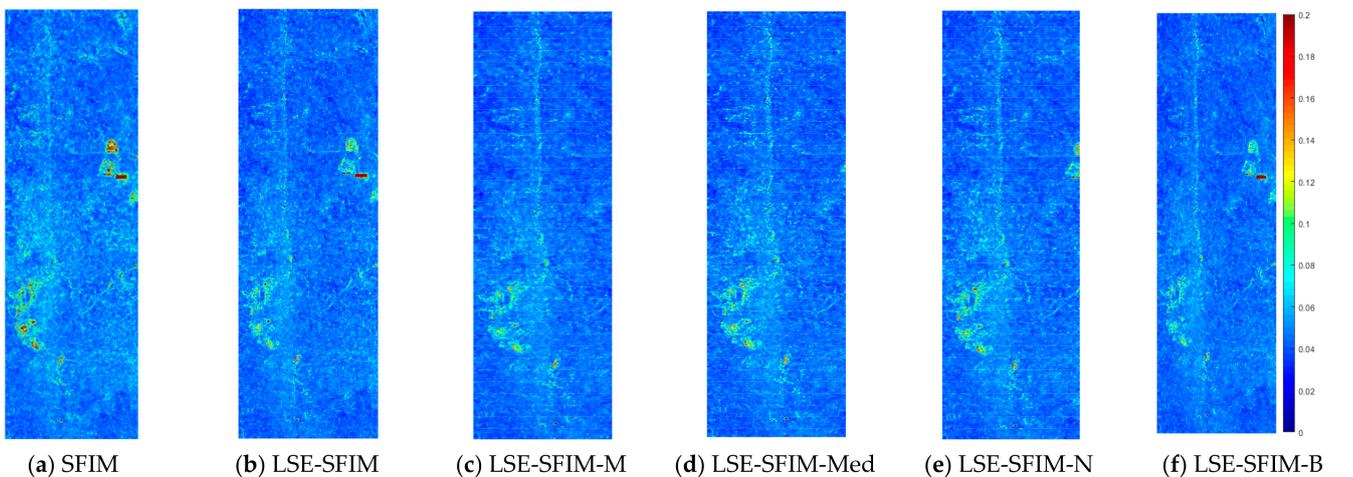


Figure 15. SAM map of six algorithms for HyMap Rodalquilar data experiment results.

3.2.4. Influence of Spatial Scale Factor between MSI and HSI

The above experiments have proved the LSE-SFIM-B algorithm to be effective among all SFIM-based algorithms. It would be interesting to know the performance of the proposed LSE-SFIM-B algorithm according to the spatial scale factor between the high-resolution MSI and low-resolution HSI images. In order to see how the algorithm performs on different scale factors, the Pavia University data are used in this section, where the spatial scale factors (SF) are set to SF = 2, 4, 8, respectively. The performance comparison is given in Table 5, where the spatial scale value is the down-sampling rate of the HSI data.

Table 5. Performance of different spatial scale factors of MSI and HSI data using LSE-SFIM-B algorithm for Pavia University.

Indexes/Spatial Scales	2	4	8
PSNR	44.1518	43.2708	42.1976
SAM	2.1803	2.4104	2.6762
CC	0.99553	0.99458	0.99362
$Q2^n$	0.93804	0.93213	0.8975
RMSE	0.005912	0.006572	0.007333
ERGAS	2.5732	1.4025	0.76253

The results in Table 5 are very interesting and show that the PSNR, SAM, CC, $Q2^n$ and RMSE values tend to worsen as the spatial scale factor increases, while the ERGAS value becomes smaller (better) as the spatial scale factor increases.

3.3. Performance Analysis of the Proposed SFIM-Based Algorithm and Other Commonly Used Algorithms

In order to compare the fusion performance of the proposed SFIM-based algorithm with the existing representative algorithms, this section chooses some state-of-the-art algorithms for comparison. Three state-of-the-art algorithms have been used in this section, which are CNMF proposed in 2012, HySure proposed in 2015, and FUSE proposed in 2018. Since the LSE-SFIM-B method has been proved to be most effective in the previous section, we will use this one for comparison. Since the robustness of the algorithm for different data sets has been proved in the previous section, in this section only the Chikusei data set is used to reduce repetition.

The experimental settings are as follows: (1) for the Chikusei data set, the number of endmembers (D) is set to $D = 30$ for any algorithm needed. (2) In CNMF algorithm, the maximum number of iterations for inner loops (I_{in}) and the maximum number of iterations for outer loops (I_{out}) are $I_{in} = 200$, $I_{out} = 2$. (3) In the HySure algorithm, the parameters are set to $\lambda_\varphi = 10^{-3}$, $\lambda_B = \lambda_R = 10$.

Figure 16a–e are the fusion results of the Chikusei dataset through five algorithms, which are conventional SFIM, the proposed LSE-SFIM-B, CNMF, HySure, and FUSE. The visual effects seem to be similar to the last four algorithms, in which the SFIM seems to have the worst performance. In order to further evaluate the performance of the proposed LSE-SFIM-B algorithm and the other three state-of-the-art algorithms, Table 6 gives the comparison for objective evaluation indicators PSNR, SAM, CC, $Q2^n$, RMSE and ERGAS.

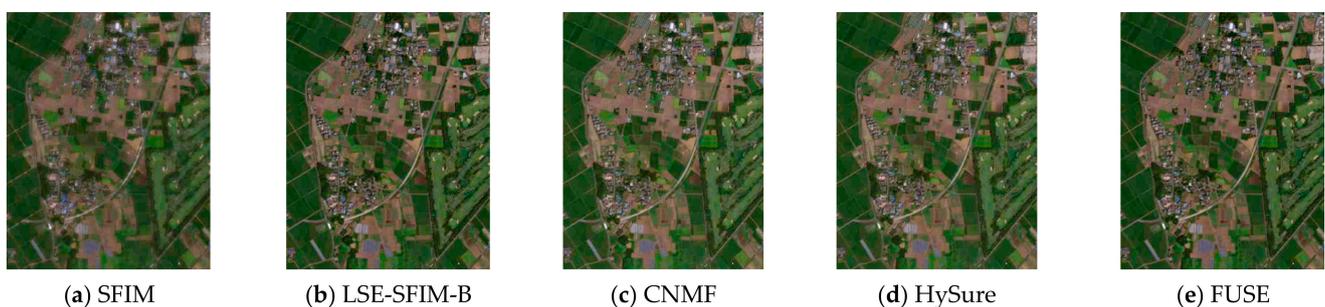
**Figure 16.** Fusion results of Chikusei dataset by five different algorithms.

Table 6. Quantitative comparison of fusion results by five different algorithms for Chikusei.

Evaluation Index	SFIM	LSE-SFIM-B	CNMF	HySure	FUSE
PSNR	24.0379	46.6653	46.1716	47.3149	45.4159
SAM	7.4477	1.3432	1.2497	1.1544	1.4699
CC	0.76329	0.99341	0.98988	0.99093	0.98855
Q2 ⁿ	0.35951	0.91992	0.9485	0.9606	0.91975
RMSE	0.4142	0.0037586	0.0035002	0.0032553	0.0044402
ERGAS	6.1140	1.2483	1.5456	1.4725	1.6222

It can be seen that the conventional SFIM algorithm has the worst performance of all six indicators. The other four algorithms, including the proposed LSE-SFIM-B, CNMF, HySure, and FUSE, have similar performance. In the HySure algorithm four of the six indicators are optimal, and the other two optimal indicators are obtained by the proposed LSE-SFIM-B algorithm. However, as mentioned in Section 1, the SFIM-based algorithm is simple to calculate and easy to implement. In order to verify the computational complexity of these different algorithms, Table 7 shows the computing time comparisons for different algorithms, and the proposed LSE-SFIM-B algorithm has the least computing time. To further show the time-efficient superiority of LSE-SFIM-B, the speed-up ratio is also provided in Table 7 which is calculated using the computing time of the other three algorithms (CNMF, HySure, and FUSE) divided by the computing time of the proposed LSE-SFIM-B algorithm. As a result, even though the HySure algorithm has four performance indicators that are better than LSE-SFIM-B algorithm, it is time consuming, especially when the data set is very large, the LSE-SFIM-B could demonstrate its excellent time efficiency while maintaining the performance.

Table 7. Computational complexity analysis by five different algorithms for Chikusei.

Algorithms	LSE-SFIM-B	CNMF	HySure	FUSE
Computing time/second	0.77	32.36	317.43	3.00
speed-up ratio/times	–	42.03	412.25	3.40

4. Discussion and Conclusions

This paper proposes a spatial-enhanced LSE-SFIM algorithm for HSI-MSI images fusion. The contributions of the proposed algorithm can be summarized as follows:

1. Improving the performance of tradition SFIM algorithm. The traditional SFIM fusion algorithm has problems such as blurred image edge information and insufficient spatial detail information. In this paper, two steps are taken to solve the above problems: (1) LSE is used to adjust the obtained simulated low-spatial MSI' so that the linear regression can minimize the spatial information error, and the simulated MSI' can have as much as possible the same spatial information as HSI; (2) four different spatial filters are then used to further improve the detailed spatial information in the process of up-sampling, and the experimental results show that the use of bilinear interpolation in LSE-SFIM-B fusion algorithm has the best performance among all SFIM based algorithms.
2. Achieving similar performance with much less computing time. This paper also employs three state-of-the art algorithms (CNMF, HySure and FUSE) to compare the performance, and the experimental results show that the proposed LSE-SFIM-B algorithm can achieve similar performance as these, while the computing time is much less. As a result, in the case of high time requirements or in the case of processing a very large data set, the proposed LSE-SFIM-B algorithm can show a good ability in both processing performance and time effect with practical significance.

The proposed algorithm can achieve a good performance in most cases, performs better than traditional SFIM algorithm with better spatial preserving and less spectral

distortion, and also has less computational complexity than the state-of-the-art fusion algorithms. However, the spectral fidelity is not good enough, since the SFIM-based model performs the fusion band by band, without considering the spectral correlations. Adding spectral constraints to the model can be considered in a future study.

Author Contributions: Conceptualization, Y.W.; Methodology, Y.W., Q.Z. and Y.S.; Experiments: Y.S.; Data Curation, Y.W. and M.S.; Formal Analysis, M.S. and C.Y.; Writing—Original Draft, Y.W. and Y.S.; Writing—Review & Editing: Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: The work of Y. Wang was supported in part by the National Nature Science Foundation of China (61801075), China Postdoctoral Science Foundation (2020M670723), the Fundamental Research Funds for the Central Universities (3132019341), and Open Research Funds of State Key Laboratory of Integrated Services Networks, Xidian University (ISN20-15). The work of Meiping Song was supported by the National Nature Science Foundation of China (61971082).

Conflicts of Interest: The authors declare no conflict of interest.

References

- McCabe, M.F.; Rodell, M.; Alsdorf, D.E.; Miralles, D.G.; Uijlenhoet, R.; Wagner, W.; Lucieer, A.; Houborg, R.; Verhoest, N.E.C.; Franz, T.E.; et al. The future of Earth observation in hydrology. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3879–3914. [[CrossRef](#)]
- Selva, D.; Krejci, D. A survey and assessment of the capabilities of Cubesats for Earth observation. *Acta Astronaut.* **2012**, *74*, 50–68. [[CrossRef](#)]
- Xiang, S.; Wang, L.; Xing, L.; Du, Y.; Zhang, Z. Knowledge-based memetic algorithm for joint task planning of multi-platform earth observation system. *Comput. Ind. Eng.* **2021**, *160*, 107559. [[CrossRef](#)]
- Shayeganpour, S.; Tangestani, M.H.; Gorsevski, P.V. Machine learning and multi-sensor data fusion for mapping lithology: A case study of Kowli-kosh area, SW Iran. *Adv. Space Res.* **2021**, *68*, 3992–4015. [[CrossRef](#)]
- Si, Y.; Lu, Q.; Zhang, X.; Hu, X.; Wang, F.; Li, L.; Gu, S. A review of advances in the retrieval of aerosol properties by remote sensing multi-angle technology. *Atmos. Environ.* **2021**, *244*, 117928. [[CrossRef](#)]
- Zhang, L.P.; Shen, H.F. Progress and future of remote sensing data fusion. *J. Remote Sens.* **2016**, *20*, 1050–1061.
- Tohid, Y.; Farhang, A.; Ali, A.; Calagari, A.A. Integrating geologic and Landsat-8 and ASTER remote sensing data for gold exploration: A case study from Zarshuran Carlin-type gold deposit, NW Iran. *Arab. J. Geosci.* **2018**, *11*, 482.
- Erdelj, M.; Natalizio, E.; Chowdhury, K.R.; Akylidiz, I.F. Help from the Sky: Leveraging UAVs for Disaster Management. *IEEE Pervasive Comput.* **2017**, *16*, 24–32. [[CrossRef](#)]
- San Juan, R.F.V.; Domingo-Santos, J.M. The Role of GIS and LIDAR as Tools for Sustainable Forest Management. *Front. Inf. Syst.* **2018**, *1*, 124–148.
- Zhu, Q.; Zhang, J.; Ding, Y.; Liu, M.; Li, Y.; Feng, B.; Miao, S.; Yang, W.; He, H.; Zhu, J. Semantics-Constrained Advantageous Information Selection of Multimodal Spatiotemporal Data for Landslide Disaster Assessment. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 68. [[CrossRef](#)]
- Khanal, S.; Fulton, J.; Shearer, S.A. An overview of current and potential applications of thermal remote sensing in precision agriculture. *Comput. Electron. Agric.* **2017**, *139*, 22–32. [[CrossRef](#)]
- Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [[CrossRef](#)]
- Zhou, L.; Chen, N.; Chen, Z.; Xing, C. ROSSC: An Efficient Remote Sensing Observation-Sharing Method Based on Cloud Computing for Soil Moisture Mapping in Precision Agriculture. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 5588–5598. [[CrossRef](#)]
- Wu, B.; Yu, B.; Yao, S.; Wu, Q.; Chen, Z.; Wu, J. A surface network based method for studying urban hierarchies by night time light remote sensing data. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 1377–1398. [[CrossRef](#)]
- Zhang, Z.; Liu, F.; Zhao, X.; Wang, X.; Shi, L.; Xu, J.; Yu, S.; Wen, Q.; Zuo, L.; Yi, L.; et al. Urban Expansion in China Based on Remote Sensing Technology: A Review. *Chin. Geogr. Sci.* **2018**, *28*, 727–743. [[CrossRef](#)]
- Shen, Q.; Yao, Y.; Li, J.; Zhang, F.; Wang, S.; Wu, Y.; Ye, H.; Zhang, B. A CIE Color Purity Algorithm to Detect Black and Odorous Water in Urban Rivers Using High-Resolution Multispectral Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6577–6590. [[CrossRef](#)]
- Carper, W.J.; Lillesand, T.M.; Kiefe, R.W. Use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 459–467.
- Chavez, P.S.; Sides, S.C.; Anderson, J.A. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Photogramm. Eng. Remote Sens.* **1991**, *57*, 265–303.
- 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](#)]

20. Sulaiman, A.G.; Elashmawi, W.H.; Eltaweel, G. IHS-based pan-sharpening technique for visual quality improvement using KPCA and enhanced SML in the NSCT domain. *Int. J. Remote Sens.* **2021**, *42*, 537–566. [[CrossRef](#)]
21. Li, Y.; Qu, J.; Dong, W.; Zheng, Y. Hyperspectral pansharpening via improved PCA approach and optimal weighted fusion strategy. *Neurocomputing* **2018**, *315*, 371–380. [[CrossRef](#)]
22. Aiazzi, B.; Alparone, L.; Arienzo, A.; Garzelli, A.; Loli, S. Fast multispectral pansharpening based on a hyper-ellipsoidal color space. In Proceedings of the Conference on Image and Signal Processing for Remote Sensing XXV, Strasbourg, France, 11 September 2019.
23. Singh, R.; Khare, A. Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach. *Inf. Fusion* **2014**, *19*, 49–60. [[CrossRef](#)]
24. Liu, J.G. Smoothing Filter-based Intensity Modulation: A spectral preserve image fusion technique for improving spatial details. *Int. J. Remote Sens.* **2000**, *21*, 3461–3472. [[CrossRef](#)]
25. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. MTF-tailored Multiscale Fusion of High-resolution MS and Pan Imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 591–596. [[CrossRef](#)]
26. Yokoya, N.; Mayumi, N.; Iwasaki, A. Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 528–537. [[CrossRef](#)]
27. 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](#)]
28. Li, S.; Dian, R.; Fang, L.; Bioucas-Dias, J. Fusing Hyperspectral and Multispectral Images via Coupled Sparse Tensor Factorization. *IEEE Trans. Image Process.* **2018**, *27*, 4118–4130. [[CrossRef](#)]
29. 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.
30. 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](#)]
31. Wei, Q.; Dobigeon, N.; Tourneret, J.-Y. Fast Fusion of Multi-Band Images Based on Solving a Sylvester Equation. *IEEE Trans. Image Process.* **2015**, *24*, 4109–4121. [[CrossRef](#)]
32. Ren, X.; Lu, L.; Chanussot, J. Toward Super-Resolution Image Construction Based on Joint Tensor Decomposition. *Remote Sens.* **2020**, *12*, 2535. [[CrossRef](#)]
33. Lu, X.; Yang, D.; Zhang, J.; Jia, F. Hyperspectral Image Super-Resolution Based on Spatial Correlation-Regularized Unmixing Convolutional Neural Network. *Remote Sens.* **2021**, *13*, 4074. [[CrossRef](#)]
34. Zhang, L.; Nie, J.; Wei, W.; Li, Y.; Zhang, Y. Deep Blind Hyperspectral Image Super-Resolution. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 2388–2400. [[CrossRef](#)] [[PubMed](#)]
35. Bedini, E.; Van Der Meer, F.; van Ruitenbeek, F. Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain. *Int. J. Remote Sens.* **2008**, *30*, 327–348. [[CrossRef](#)]