World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:11, No:9, 2017

A Hybrid Image Fusion Model for Generating High Spatial-Temporal-Spectral Resolution Data Using OLI-MODIS-Hyperion Satellite Imagery

Yongquan Zhao, Bo Huang

***Abstract*—**Spatial, Temporal, and Spectral Resolution (STSR) are three key characteristics of Earth observation satellite sensors; however, any single satellite sensor cannot provide Earth observations with high STSR simultaneously because of the hardware technology limitations of satellite sensors. On the other hand, a conflicting circumstance is that the demand for high STSR has been growing with the remote sensing application development. Although image fusion technology provides a feasible means to overcome the limitations of the current Earth observation data, the current fusion technologies cannot enhance all STSR simultaneously and provide high enough resolution improvement level. This study proposes a Hybrid Spatial-Temporal-Spectral image Fusion Model (HSTSFM) to generate synthetic satellite data with high STSR simultaneously, which blends the high spatial resolution from the panchromatic image of Landsat-8 Operational Land Imager (OLI), the high temporal resolution from the multi-spectral image of Moderate Resolution Imaging Spectroradiometer (MODIS), and the high spectral resolution from the hyper-spectral image of Hyperion to produce high STSR images. The proposed HSTSFM contains three fusion modules: (1) spatial-spectral image fusion; (2) spatial-temporal image fusion; (3) temporal-spectral image fusion. A set of test data with both phenological and land cover type changes in Beijing suburb area, China is adopted to demonstrate the performance of the proposed method. The experimental results indicate that HSTSFM can produce fused image that has good spatial and spectral fidelity to the reference image, which means it has the potential to generate synthetic data to support the studies that require high STSR satellite imagery.

***Keywords*—**Hybrid spatial-temporal-spectral fusion, high resolution synthetic imagery, least square regression, sparse representation, spectral transformation.

I. Introduction

NOWADAYS, a large number of Earth observation satellites have been launching, with various STSR [1], [2], which contribute significantly to Earth surface or atmosphere environment monitoring ability. It seems that there is a boom in Earth observation field, nevertheless, because of the limitations of satellite sensor's hardware technology and budget constraints, there exist compromises between STSR of satellite sensors. That is to say, even though so many satellites have been launched, none of them can obtain high STSR data simultaneously. These compromises limit the application of existing remotely sensed data significantly, especially for the remote sensing applications that require fine spatial details, long-term and frequent coverage, and hyper-spectral (HS) satellite imagery, such as precise global or regional change detection, urban dynamic monitoring, natural disaster monitoring, real-time air quality monitoring, etc.

As a low-cost and flexible solution to overcome the compromises between different types of resolution of Earth observation satellites, image fusion can produce synthetic high-resolution satellite images [3], [4]. However, most of the existing image fusion methods mainly focus on part types of the resolution enhancement issues, e.g. spatial and spectral fusion [5]-[7], spatial and temporal fusion [8]-[10], which cannot generate synthetic images with high STSR simultaneously. Although some image fusion methods took the three types of resolution into consideration simultaneously [11], [12], they just adopted MODIS image as both high spectral and temporal resolution data source. Nevertheless, the spectral resolution of MODIS is still too limited to be called as HS image. The SPATS method proposed in [13] blended high spectral resolution from HS image, like Hyperion, but the spatial resolution of the final fused image was also limited (which did not consider the high spatial resolution of Landsat panchromatic image). Meanwhile, its temporal change prediction ability is also limited to phenological change. Hence, a comprehensive image fusion algorithm that can blend STSR from different satellite sensors together is still a big challenge for the current image fusion algorithms.

This study presents a HSTSFM to achieve STSR enhancement together. The hybrid image fusion model contains spatial-spectral image fusion, spatial-temporal image fusion, and temporal-spectral image fusion, which are conducted based on different fusion schemes. It can blend high-spatial, high-temporal, and high-spectral resolution from multi-sensors together to produce high STSR satellite images, which aims to satisfy the growing demand of high STSR satellite images in various remote sensing applications. The remainder of this paper is organized as follows. The multi-source experimental data description is given in Section II. The image fusion sub-modules of HSTSFM is introduced in Section III. The experimental results, the fusion precision assessment criteria, and the fusion result analysis are shown in Section IV. Finally, the conclusions are presented in Section V.

II. Experimental Dataset

Three data sources are adopted as desired STSR sources,

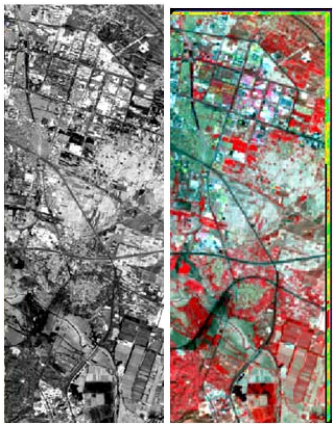
Yongquan Zhao is with the Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong, China (phone: 852-6350-7498; e-mail: yqzhao@link.cuhk.edu.hk).

Bo Huang is with the Department of Geography and Resource Management and the Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China (e-mail: bohuang@cuhk.edu.hk).

which are the Landsat-8 OLI, MODIS, and Hyperion images. The fused synthetic high STSR image obtains high spatial resolution from OLI panchromatic (PAN) image, i.e. 15 m, high temporal resolution from MODIS, i.e. daily observations, and high spectral resolution from Hyperion, i.e. hundreds of spectral bands. The fusion target is to predict the HS image on a prediction date with 15 m spatial resolution from a prior date. Meanwhile, we chose the sub-urban area of Beijing, China as our test region, which involves both distinct phenological and land cover type changes. All the test data are exempted from atmospheric correction because the PAN images are involved in the image fusion procedure, which are the Top-Of-Atmosphere (TOA) radiance images. In addition, all of the experimental images were projected to the UTM projection under the WGS-84 datum. Meanwhile, all of the test images were up-sampled to 15 m spatial resolution by using the bi-cubic interpolation method for the afterwards image fusion. Small spatial offsets may exist among them after the re-projection and re-sampling, hence, these images were geometrically co-registered by applying an optimal offset [14].

A. Landsat-8 OLI Data

Landsat-8 OLI captures 8 multi-spectral (MS) bands with 30 m spatial resolution and 1 PAN image with 15 m spatial resolution every 16 days. Hence, 6 MS bands that have counterparts among the MODIS sensor bands are selected out, i.e., blue, green, red, near-infrared (NIR), short wave infrared 1 (SWIR1), and short wave infrared 2 (SWIR2). The PAN and MS images captured on May 25, 2015 are adopted as prior images, which has the image size of 672 × 256 pixels for 15 m spatial resolution. Fig. 1 shows the OLI PAN and MS (3D spectral cube) images, it should be noted that the MS image is displayed with NIR, red, and green bands as RGB composite images.

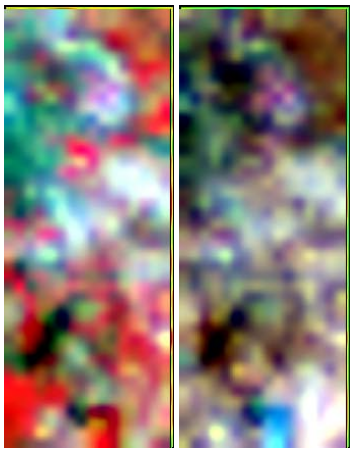


(a) (b)

Fig. 1 OLI images captured on May 25, 2015 (a) PAN image, (b) MS image

B. MODIS Data

The MODIS images in this study were aggregated from the corresponding OLI images because the radiometric and geometric inconsistencies can negatively affect the accuracy of spatiotemporal image fusion results [14], [15]. Using simulated MODIS images can exclude these negative factors' interference and can demonstrate the effects of the spatiotemporal image fusion explicitly. The NIR and red bands were aggregated to 250 m, and the other four bands were aggregated to 500 m to simulate the spatial resolution of MODIS sensor. Beside the base date (T_1) of the OLI data, i.e. May 25, 2015, another time point on February 21, 2016 is adopted as the prediction date (T_2) in the spatiotemporal image fusion module. These two dates located in the end of spring and winter, respectively. During that time period, distinct phenological changes and land cover type changes existed simultaneously. Fig. 2 shows the MODIS (3D spectral cube) images on the base and prediction dates, and they are both displayed with NIR, red, and green bands as RGB composite images.



(a) (b)

Fig. 2 Simulated MODIS images (a) May 25, 2015, (b) February 21, 2016

C. Hyperion Data

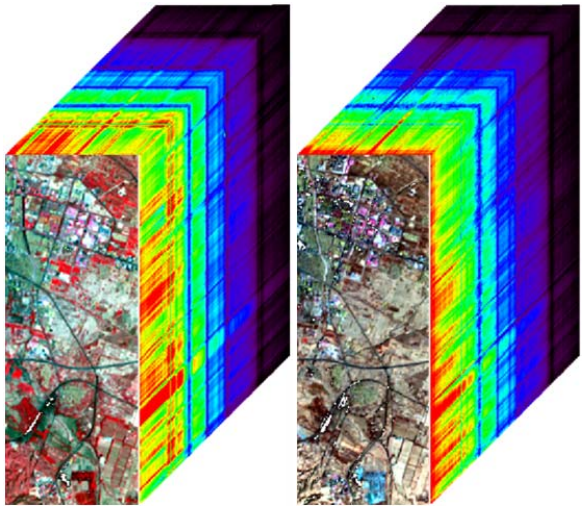
Because the temporal resolution (200 days) of Hyperion is too low to obtain their counterpart images with MODIS sensor, we adopted the rare available Hyperion images that were obtained on May 20, 2015 and February 26, 2016 to act as the corresponding counterparts with the simulated MODIS images. Since their capture dates are very close, it is reasonable to use the Hyperion images on these two dates as substitutes. The Hyperion sensor captures 242 bands that are located in the spectral range from 0.4 µm to 2.5 µm with a 30 m spatial resolution. Because of the calibration absence, noise disturbance, water vapor absorption, and vertical stripes effects, subsets of 146 good-quality bands out of 242 bands were selected for our study, see Table I.

During the algorithm experiments, the Hyperion image on May 20, 2015 will be used as prior HS image; the one on February 26, 2016 will be employed as spectral reference image to demonstrate the spectral fidelity of the fused image. The two Hyperion image cubes are shown in Fig. 3, it should be noted that the Hyperion image is displayed with the selected band index 38, 20, and 10 (corresponding to the NIR, red, and

green bands of OLI) as RGB composite images.

TABLE I
SELECTED HYPERION BAND SEQUENCE

Selected band index	Original band index	Spectral range (nm)
1-47	11-57	457-925
48-66	79-97	932-1114
67-85	101-119	1154-1336
86-115	135-164	1497-1790
116-146	188-218	2032-2335



(a) (b)

Fig. 3 Hyperion images. (a) May 20, 2015, (b) February 26, 2016

The overall fusion process that uses the three data sources mentioned above is shown in Fig. 4.

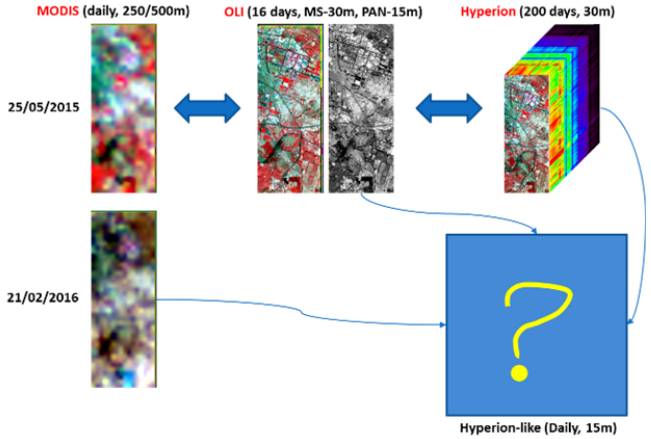


Fig. 4 The fusion of STSR from multi-source sensors

III. METHODOLOGY

The hybrid image fusion workflow is shown in Fig. 5, which comprises three image fusion modules: (1) spatial-spectral image fusion to blend OLI-PAN and OLI-MS/Hyperion images for the following steps; (2) spatial-temporal image fusion to predict the 15 m MS image on the prediction date; (3) temporal-spectral image fusion to produce the 15 m HS image on the prediction date. In the end, the precision of the hybrid image fusion model is evaluated based on the fused 15 m HS image on the prediction date from step 1.

A. Spatial-Spectral Image Fusion

Spatial-spectral image fusion usually includes PAN/MS fusion, PAN/HS fusion, MS/HS fusion [11], this study mainly focuses on PAN/MS and PAN/HS fusion, which can be also called as pan-sharpening. An Enhanced Synthetic Variable Ratio (ESVR) algorithm is developed to obtain high spatial resolution OLI-MS and Hyperion images at T_1, which is based on the Synthetic Variable Ratio (SVR) pan-sharpening method [16]. The SVR method is formulated as:

$$Fused\ MS_i = MS_i \times (PAN/PAN_{Syn}) \quad (1)$$

where i is the MS band index, PAN and MS are the input panchromatic and multi-spectral images, $Fused\ MS$ is the pan-sharpened multi-spectral image, and PAN_{Syn} is a synthetic panchromatic image that has the same spatial resolution with the input multi-spectral image, which is produced by:

$$PAN_{Syn} = \sum \varphi_i MS_i \quad (2)$$

The coefficient φ_i is calculated by,

$$PAN = \sum \varphi_i MS_i \quad (3)$$

which adopts a multiple linear regression between the input panchromatic and multi-spectral images. In addition, there is a histogram standardization process before the image fusion steps stated in (1), (2), and (3) to reduce spectral distortion in spatial-spectral fusion result, which conduct a histogram transformation upon the input PAN and MS images to make them have the same mean and standard deviation (Stdev). However, since the common mean and Stdev are determined by human subjectively, the histogram standardization will affect the fused image's mean and Stdev, which may have large bias with the real mean and Stdev. On that account, a histogram de-standardization is performed on the SVR fusion result, which uses predicted Stdev values of the high spatial resolution MS bands and mean values of the input MS bands to conduct an inverse process of the previous histogram standardization to make the fused bands have proper mean and Stedv values. The reason why we adopt the input MS bands' mean values directly but predict the Stdev values is that the MS bands with different spatial resolution have nearly the same mean but different Stdev values.

As to the prediction of the Stdev values of high spatial resolution MS image bands, it is achieved by a multi-order polynomial fitting method. An iteration process is used to determine the optimal fitting order between 1 and 5. We down-sample the input MS image into an image series with a series of diminishing spatial resolution, see the examples in Fig. 6, and the down-sampled scales and its corresponding Stdev values are employed as independent and dependent variables respectively in the polynomial fitting process. After the relationship between the spatial resolution and the

corresponding Stdev of the MS image series is obtained, we can predict the Stdev of the MS image with higher spatial resolution than the input MS image. It should be noted that the PAN/MS and PAN/HS fusion are both achieved by ESVR method in this study. Furthermore, the MS or HS bands that are used in (2) and (3) are selected based on the spectral range of OLI-PAN image, which means that the spectral ranges of the MS or HS bands in (2) and (3) are covered by the panchromatic spectral range of OLI.

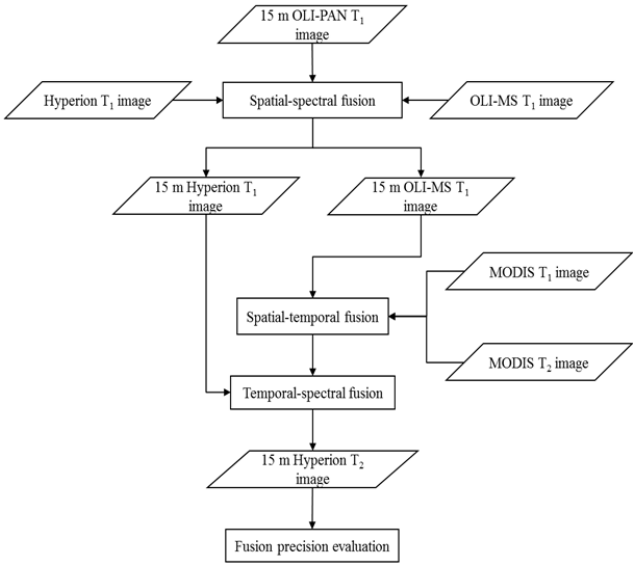


Fig. 5 Workflow of the HSTSFM

Fig. 6 Down-sampled image series with diminishing spatial resolution

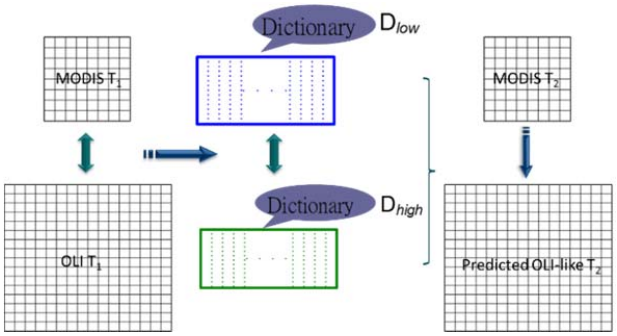


Fig. 7 One-pair image learning spatiotemporal fusion process

B. Spatial-Temporal Image Fusion

Considering that there exist both phenological and land cover type changes between the two MODIS images, and there is only one prior date's MODIS and OLI images available, the one-pair image dictionary learning method (one-pair image learning) [9] is adopted to obtain high spatial resolution OLI-MS image at T_2 based on the fused 15 m OLI-MS image at T_1 and the two MODIS images in Fig. 2. The spatiotemporal fusion process is shown in Fig. 7, D_{low} and D_{high} represent the low- and high- spatial resolution dictionaries, respectively. For more details about the one-pair learning spatiotemporal image fusion method, we refer the readers to [9].

C. Temporal-Spectral Image Fusion

After the 15 m OLI-MS and Hyperion images at T_1 are obtained by the ESVR method, the 15 m OLI-MS image at T_2 is predicted by the one-pair image learning method. The prior HS image at T_1 has the same spatial resolution with the two prior MS images, hence temporal-spectral image fusion is conducted here to produce the 15 m Hyperion image at T_2. The Color Resolution Improvement Software Package (CRISP) algorithm [6] is modified in this study to obtain the 15 m Hyperion image at T_2, i.e. the final predicted high STSR image. Since there is no spatial resolution difference among these input images, the image sharpening process based on Butterworth filter in the original CRISP can be removed in this circumstance. Additionally, we assume that the spectral relationship between the MS and HS image at T_1 can be applied to T_2, hence the 15 m Hyperion image at T_2 can be calculated by,

$$\widehat{HS}_{T_2} = G \times MS_{T_2} \tag{4}$$

where $\widehat{HS}_{T_2}$ is the final fused image with 15 m spatial resolution and 146 spectral bands at T_2, MS_{T_2} is the predicted 15 m OLI-MS image at T_2, and the spectral transformation matrix G is derived by a least square regression as,

$$G = HS_{T_1} \times {MS_{T_1}}^T \times \left(MS_{T_1} \times {MS_{T_1}}^T\right)^{-1} \tag{5}$$

where HS_{T_1} and MS_{T_1} are the 15 m OLI-MS and Hyperion images respectively at T_1. The framework of the temporal-spectral image fusion is shown in Fig. 8.

IV. Experimental Results and Analysis

To demonstrate the effect of the proposed HSTSFM, we compared the final fused image with the fused 15 m Hyperion image at T_2 visually and quantitatively, which is obtained by ESVR using the actual PAN and Hyperion images at T_2 because there is no such image at T_2. The predicted and actual 15 m Hyperion image at T_2 are both shown in Fig. 9, which are displayed with the selected band index 38, 20, and 10 as RGB composite images. It can be seen that both phenological and land cover type changes are captured, and the hyper-spectral information of the predicted image is also quite close to the reference image. Additionally, the spatial details from the OLI-PAN image are injected into the fusion result precisely. Meanwhile, a quantitative evaluation in terms of Correlation Coefficient (CC), Root Mean Square Error (RMSE), and Structure Similarity index (SSIM) between the two images in Fig. 9 is also performed. The evaluation criteria are: (1) the closer the CC or SSIM to one, the better; (2) the smaller and closer the RMSE to zero, the better. In this study, because there

is a temporal gap between the two acquisition dates of the input OLI and Hyperion images, we use the correlation between the pan-sharpened 15 m Hyperion image at T_1 with the reference image at T_2 as a baseline to demonstrate the effect of the hybrid image fusion model. Considering that the HS image in our study has up to 146 bands, the quantitative indices of these bands are averaged to be shown in Table II. It can be seen that the correlation is improved significantly after the fusion, which means that the proposed hybrid image fusion model can blend the high STSR from multi-sensors' imagery well.

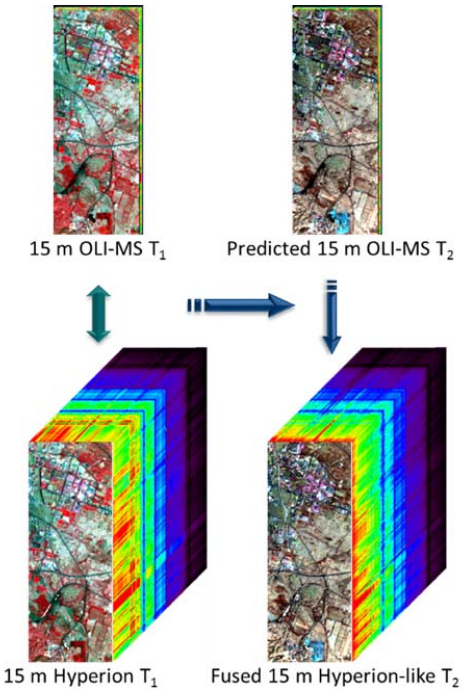
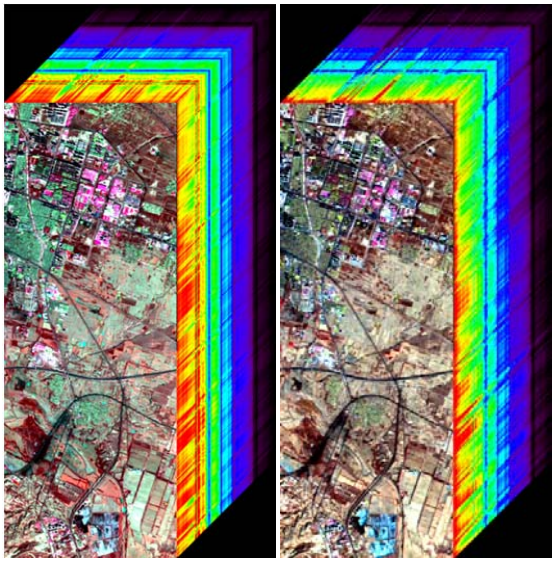


Fig. 8 Temporal-spectral image fusion process



(a) (b)

Fig. 9 The predicted (a) and actual (b) 15 m Hyperion images at T_2

TABLE II
QUANTITATIVE INDEX COMPARISON WITH REFERENCE HS IMAGE

Subject HS image	Mean of CC	Mean of RMSE	Mean of SSIM
Input T_1	0.5458	16.7321	0.3861
Predicted T_2	**0.9063**	**9.1097**	**0. 6425**

V. CONCLUSIONS

This study presents a HSTSFM to fuse high STSR from multi-source satellite imagery, which can generate high STSR images. Not only the HSTSFM can inject high resolution spatial details from PAN image into MS or HS image, but also it can capture the phenological and land cover type changes simultaneously. In addition, the temporal-spectral image fusion module in HSTSFM can enhance the spectral resolution of a MS image on a prediction date based on the prior MS-HS image pair on a base date. Meanwhile, HSTSFM is of relatively low computing complexity, which has significant practical application potential to be employed by remote sensing applications that need high STSR satellite imagery.

ACKNOWLEDGMENT

This study was supported by the Natural Science Foundation of China under Grant 41371417. We also would like to thank Dr. Huihui Song for providing the figure of the workflow of the one-pair image learning method.

REFERENCES

[1] Tong, X., Zhao, W., Xing, J., & Fu, W. (2016, July). Status and development of China High-Resolution Earth Observation System and application. In Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International (pp. 3738-3741). IEEE.
[2] The future missions of NASA's Earth Observing System, https://eospso.nasa.gov/future-missions.
[3] Zhang, Y. (2004). Understanding image fusion. Photogrammetric engineering and remote sensing, 70(6), 657-661.
[4] Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. Geoscience and Remote Sensing, IEEE Transactions on, 44(8), 2207-2218.
[5] Zhang, Y. (1999). A new merging method and its spectral and spatial effects. International Journal of Remote Sensing, 20(10), 2003-2014.
[6] Winter, M. E., Winter, E. M., Beaven, S. G., & Ratkowski, A. J. (2007, March). Hyperspectral image sharpening using multispectral data. In Aerospace Conference, 2007 IEEE (pp. 1-9). IEEE.
[7] Song, H., Huang, B., Zhang, K., & Zhang, H. (2014). Spatio-spectral fusion of satellite images based on dictionary-pair learning. Information Fusion, 18, 148-160.
[8] Huang, B., & Song, H. (2012). Spatiotemporal reflectance fusion via sparse representation. Geoscience and Remote Sensing, IEEE Transactions on, 50(10), 3707-3716.
[9] Song, H., & Huang, B. (2013). Spatiotemporal satellite image fusion through one-pair image learning. IEEE Transactions on Geoscience and Remote Sensing, 51(4), 1883-1896.
[10] Zhao, Y., & Huang, B. (2016, August). A Two-step Spatio-Temporal satellite image Fusion Model for temporal changes of various LULC under one-pair prior images scenario. In Signal Processing, Communications and Computing (ICSPCC), 2016 IEEE International Conference on (pp. 1-5). IEEE.
[11] Shen, H., Meng, X., & Zhang, L. (2016). An Integrated Framework for the Spatio–Temporal–Spectral Fusion of Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing, 54(12), 7135-7148.
[12] Huang, B., Zhang, H., Song, H., Wang, J., & Song, C. (2013). Unified fusion of remote-sensing imagery: generating simultaneously high-resolution synthetic spatial–temporal–spectral earth observations. Remote sensing letters, 4(6), 561-569.

[13] Zhang, L., Fu, D., Sun, X., Chen, H., & She, X. (2016, April). A spatial-temporal-spectral blending model using satellite images. In IOP Conference Series: Earth and Environmental Science (Vol. 34, No. 1, p. 012042). IOP Publishing.
[14] Gevaert, C. M., & García-Haro, F. J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. Remote sensing of Environment, 156, 34-44.
[15] Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., & Lefsky, M. A. (2016). A flexible spatiotemporal method for fusing satellite images with different resolutions. Remote Sensing of Environment, 172, 165-177.
[16] Zhang, Y. (2008). U.S. Patent No. 7,340,099. Washington, DC: U.S. Patent and Trademark Office.