

Simulation of Low-Resolution Panchromatic Images by Multivariate Linear Regression for Pan-Sharpener IKONOS Imageries

Zhongwu Wang, Shunxi Liu, Shucheng You, and Xin Huang

Abstract—The extraction of spatial details is crucial for fusion quality. An efficient way is to exploit the difference between high-resolution panchromatic (Pan) images and low-resolution Pan (LRP), which is to be simulated by weighted average value from low-resolution multispectral images. To obtain the weighting coefficients with multivariate linear regression, three issues were discussed, and corresponding solutions were proposed in this letter. The proposed method consists of separating high-frequency pixels from low-frequency pixels using support vector machine and selecting observations that are evenly distributed by a bucketing technique and forcing coefficients to be sound physically by constrained least squares. Validation experiments are undertaken using three IKONOS data sets, and fusion results are compared against four popular methods. The results show that the proposed method can simulate LRP soundly and therefore achieve a better fusion quality.

Index Terms—Bucketing technique, constrained least squares, fusion, multivariate linear regression, support vector machine (SVM).

I. INTRODUCTION

THE OBJECTIVE of image fusion is to merge the spatial information of a high-resolution panchromatic (Pan) (HRP) image and the spectral information of a low-resolution multispectral (MS) (LRM) image to produce a high-resolution MS image (HRM). For image fusion of very high resolution data, particularly IKONOS MS+Pan images, fusion quality troubles quite a few traditional methods.

Various research works have been undertaken to confirm that spatial detail and its injecting manner are the main reasons for color distortion [1]–[3]. To extract spatial detail, one of the widely used techniques is to exploit the difference between HRP and low-resolution Pan (LRP), which is to be simulated by weighted average value of the MS bands [4]. Considering spectral similarity, the near-infrared band was added into the average [3]. Moreover, spectral response was considered to ob-

tain the weighting coefficients [5]. However, it only considered nominal spectral response in the laboratory and ignored the effect of the on-orbit operational conditions such as variability of the observed scene [2]. In addition, statistical methods, integrated into the estimation of the coefficients, e.g., multivariate linear regression [2], have been approved to achieve sorts of satisfied results. The multivariate linear regression technique does not need to consider about a prior knowledge, and therefore, it can be expected to be useful.

The remainder of this letter is organized as follows. The Pan-sharpening algorithm using multivariate linear regression and its problems are presented in Section II. The improved method is outlined in Section III. Experiments and conclusions are shown in Sections IV and V.

II. PAN-SHARPENING WITH MULTIVARIATE LINEAR REGRESSION

The framework of image fusion [4] is described as

HRM_i = LRM_i + W_iδ, i = 1, ..., n (1)

where n is the number of MS channels, i denotes the i-th band of a multichannel image, HRM is the fusion result, LRM is an up-sampled image of MS of the same size as Pan, W denotes the injecting parameter, and δ denotes a spatial detail of redundant information. Traditionally, δ indicates the difference between HRP and LRP, while HRP is the original Pan or its transformation by a histogram matching LRP.

The simulation of LRP by multivariate linear regression from LRM can be described as [6]

LRP = ∑_{i=1}ⁿ φ_iLRM_i (2)

where φ_i is the regression coefficient.

Generally, spatial detail is high-frequency information, such as edges and noise. It would cause the main difference between LRP and HRP, while low-frequency information remains unchanged. Equation (2) can be seen as a multivariate linear regression problem and can be solved by the least squares method with observations from HRP^l, low-frequency information in HRP

HRP^l = ∑_{i=1}ⁿ φ_iLRM_i. (3)

Digital Object Identifier 10.1109/LGRS.2010.2040706

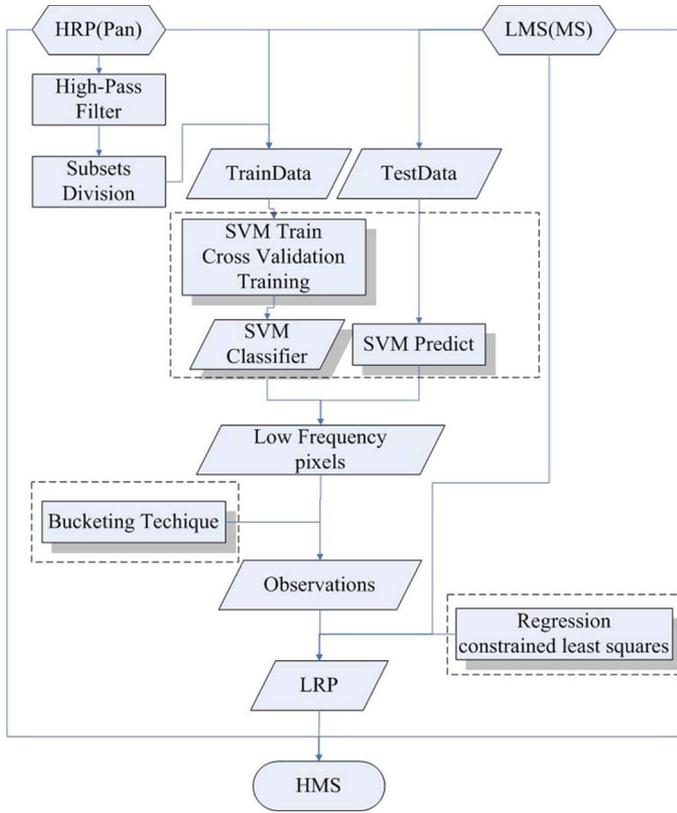


Fig. 1. Flow chart of the proposed method.

Nevertheless, there are a few issues to be considered in the implementation.

- 1) As the threshold of separating high-frequency pixels from low-frequency pixels is crucial for the simulation accuracy, it is desirable to develop an adaptive method for threshold selection.
- 2) Generally, it is quite time consuming to use all the pixels from images of huge size in calculating regression coefficients. Meanwhile, all parts of the studied scene should be considered equally. Therefore, a sampling strategy is necessary to select pixels that are evenly distributed as observations.
- 3) Furthermore, prior knowledge indicates that all bands of MS having overlap in the spectral range with that of Pan are obligated to contribute positively or a minimum amount to LRP, at least. Therefore, φ_i must be limited to a defined numerical range.

III. IMPROVED WORKFLOW FOR SIMULATING LRP

Based on the support vector machine (SVM), an improved method, combined with bucketing technique and constrained least squares, is proposed to simulate LRP automatically and is shown in Fig. 1.

A. Separating High- and Low-Frequency Pixels Using SVM

The procedure of separating high- and low-frequency pixels could be seen as a classification problem. First, Laplacian high-pass filtering, which is the same as the “high-pass filter Resolution Merge” of ERDAS 9.1, is implemented on HRP with a kernel size of $2R + 1$ (where R is the spatial resolution

ratio between MS and Pan), and the resulting image is named as HP. Then, HP is divided into $p \times q$ subsets (p and q are the number of rows and columns) with positions of the maximum and minimum values in every subset found. Furthermore, the vectors, composed of gray values in the same places on MS and Pan (e.g., $[Ms_1, \dots, Ms_n, Pan]$, with n as the number of MS bands) are set to be the attributions of *traindata* as samples, and the corresponding group -1 and $+1$ represent catalogs of high- and low-frequency pixels individually in

$$\text{trainData} = \begin{bmatrix} Ms_{11}, \dots, Ms_{n1}, Pan_1 \\ Ms'_{11}, \dots, Ms'_{n1}, Pan'_1 \\ \dots \\ Ms_{1(p \times q)}, \dots, Ms_{n(p \times q)}, Pan_{(p \times q)} \\ Ms'_{1(p \times q)}, \dots, Ms'_{n(p \times q)}, Pan'_{(p \times q)} \end{bmatrix}$$

$$\text{Group} = \begin{bmatrix} -1 \\ +1 \\ \dots \\ -1 \\ +1 \end{bmatrix} \quad (4)$$

where Ms_{ij} is the gray value of the i th MS band located at the same row and column with the pixel having the maximum value in subset j on HP, and Pan_j is the gray value of Pan in the same position; Ms'_{ij} and Pan'_j are those having the minimum value.

Afterward, the attributions of *testdata* are defined to be all the vectors composed of gray values of MS and Pan, having the same format with *traindata*; refer to

$$\text{testData} = \begin{bmatrix} Ms_{11}, \dots, Ms_{n1}, Pan_1 \\ \dots \\ Ms_{1(row \times col)}, \dots, Ms_{n(row \times col)}, Pan_{(row \times col)} \end{bmatrix} \quad (5)$$

where *row* and *col* denote the height and width of Pan.

Finally, SVM is used to train *traindata*, followed by the classification of *testdata*. Considering that the number of attributes is equal to the number of MS bands (e.g., four for IKONOS), which is far less than the number of samples, the radial basis function kernel is recommended to be the kernel model [7]. Cross validation is used to reduce the effect of wrong samples locally as well.

B. Selecting Pixels Using Bucketing Technique

Bucketing technique can randomly select points that are evenly distributed by assuming that each bucket has the same number of points. However, in our case, the amounts of low-frequency pixels are not the same among buckets. To increase the probability of the bucket having more low-frequency pixels, the range $[0, 1]$ is divided into $p \times q$ buckets (corresponding to $p \times q$ subsets) so that the length of the i th bucket is $l_i / \sum l_i$, where l_i is the number of low-frequency pixels attached to the i th bucket [8].

As shown in Fig. 2(a), the HP is divided into 2×2 buckets with nine pixels in each, and the shadow represents high-frequency pixel while the blank is the low-frequency one. From Fig. 2(b), the ranges of the four buckets are shown in the last column individually. Then, the improved bucketing technique is as follows.

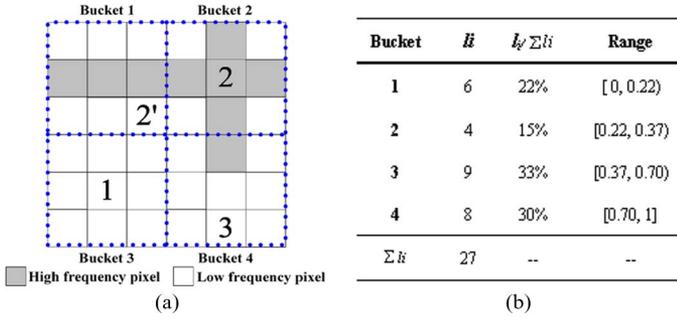


Fig. 2. Demonstration of improved bucketing technique. (a) Demonstration. (b) Range of buckets.

First, a random number (e.g., 0.5) in [0, 1] is generated, and bucket 3 is chosen since 0.5 is in its range. Second, a new random integer in [1] and [9], which indicates the serial number of a pixel to be selected in the chosen bucket, is generated. For example, an integer 5 represents pixel 1 in bucket 3 by rearranging the pixels in bucket 3 row-by-row. As pixel 1 is a low-frequency pixel, it is selected as observation and is excluded from the next selection. Otherwise, a new bucket selection and low-frequency pixel selection are done as far as pixel 2 is concerned; if the high-frequency pixel 2 is wrongly selected at first, the reselection goes again until a desired pixel 2' is selected. Then, the steps are repeated until enough pixels are selected.

C. Calculating Regression Coefficients Using Constrained Least Squares

For the problem described in (3), it is known that each band of MS having overlap in spectral range with that of Pan should contribute to LRP at a certain extent. Therefore, regression coefficients must be limited to a given numerical range, e.g., $[xMin, xMax]$, indicating the minimum and maximum contribution amount. Then, the solution is (6) if no other conditions are imposed, i.e.,

$$\min_{\varphi_i} \left\{ \left\| \sum_{i=1}^n (\varphi_i * LRM_i) - HRP \right\|^2 \right\} \text{s.t. } \begin{matrix} xMin \leq \varphi_i \leq xMax \\ i = 1, \dots, n \end{matrix} \quad (6)$$

As the recommended φ_i is 0.2 for IKONOS MS+Pan in the ‘‘Modified intensity–hue–saturation (IHS) Resolution Merge’’ of ERDAS 9.1, the default $xMin$ and $xMax$ are set to be 0.05 and 1 for simple comparison.

IV. EXPERIMENTS AND ANALYSES

Experiments are carried out using two data sets of IKONOS. The first scene [Fig. 3(a) and (b)] shows a natural area, including bare rocks on mountain areas, residential areas, terrace, and alleys among mountain areas. The second scene [Fig. 3(c) and (d)] covers ground objects including small lakes, grass and trees, blocks, and buildings in urban areas. All data sets are small fragments with $1200 * 1200$ pixels on Pan.

SVM with cross validation is tested on the two data sets. As far as data set 1 is concerned, the cross-validation rate is 90.67%. It indicates that there are nearly 10% of samples in *traindata* which failed in getting through the cross-validation procedure, and that the local minimum or maximum value in a few subsets, e.g., the hillside, will be classified incorrectly

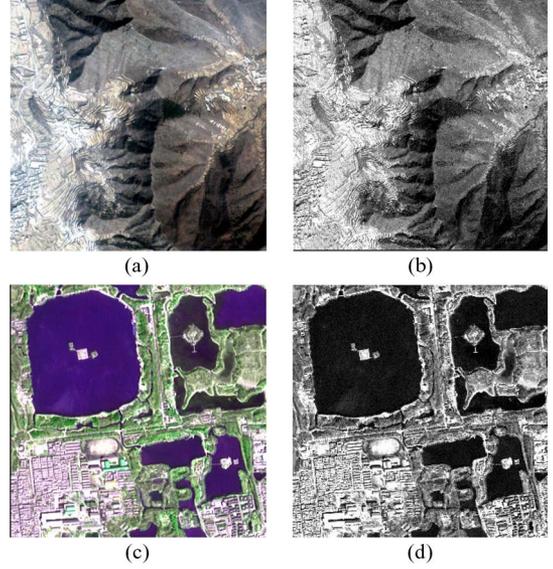


Fig. 3. Original MS and Pan of experiment data sets. (a) and (b) MS and Pan of data set 1. (c) and (d) MS and Pan of data set 2.

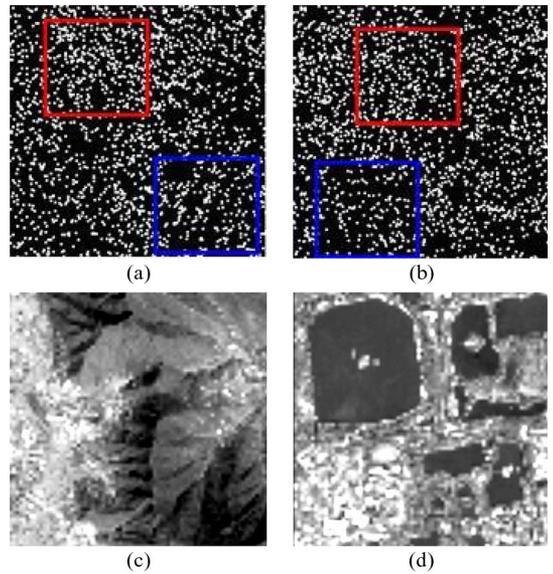


Fig. 4. Selected pixels and simulated LRP. (a) and (b) Selected pixels of data sets 1 and 2. (c) Simulated LRP of data sets 1 and 2.

without cross validation. On data set 2, the phenomenon is magnified, while the cross-validation rate is 88.67%. The false classifications highly occur in lake areas on the top left and top right of the image. Fig. 4(a) shows the efficiency of the improved bucketing technique on selecting 2000 evenly distributed pixels with low frequency and the high probability for the selection of pixels in buckets with rare high-frequency pixels in data set 1, such as the large amount of selected pixels in red rectangle versus the quite few of those in blue rectangle. A similar situation is encountered in Fig. 4(b) for data set 2 as well. Matching the red rectangle to the original MS, the pixels of the hillside on data set 1 and lakes on data set 2 get high chances.

Then, a constrained least squares method is used to limit φ_i . Fig. 5(c) and (d) shows the simulated LRPs of the two data sets. While comparing LRP with the corresponding Pan visually, on data set 1, digital numbers of the terrace tend to be high, those of

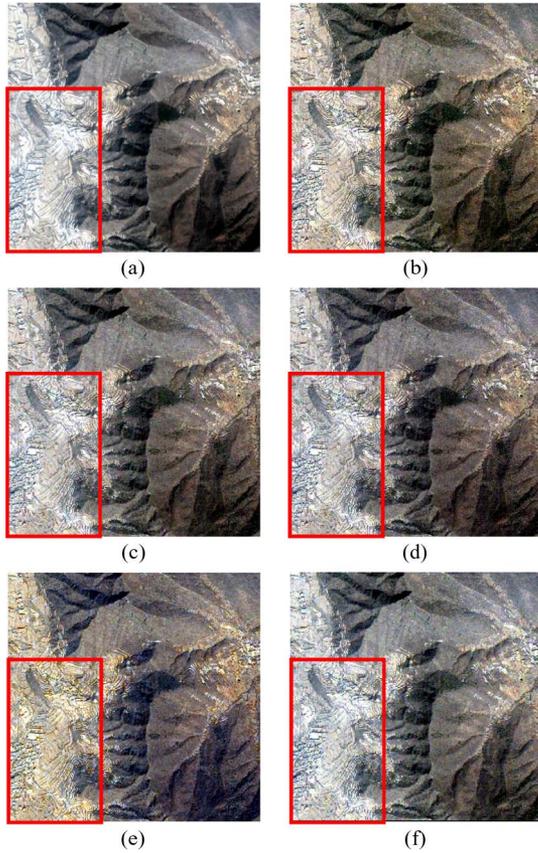


Fig. 5. Fusion experiments on data set 1. (a) Reference. (b) GS_AVE. (c) GS_LS. (d) GS_CLS. (e) FIHSSA. (f) PANSHARP.

TABLE I
REGRESSION COEFFICIENTS BY SIMULATION METHODS

Data set	Simulating method	Regression coefficients
1	Average	0.25, 0.25, 0.25, 0.25
	Least Squares	0.30, -0.07 , 0.48, 0.20
	Constrained least squares	0.15, 0.27, 0.08, 0.37
2	Average	0.25, 0.25, 0.25, 0.25
	Least Squares	-0.21 , 0.62, 0.01, 0.56
	Constrained least squares	0.05, 0.39, 0.05, 0.42

the residential areas are a little lower, and the black areas are the shadow of mountains on both images. On data set 2, buildings and blocks seem to be white, grass and trees are darker, and the lakes are black. Generally, LRPs inherit most low frequency from HRP for all data sets.

To indicate the superiority of the proposed method, comparisons are undertaken among average, least squares, and constrained least squares (see Table I). The following can be noticed.

- 1) The least squares method obtains negative coefficients on both data sets compared with the constrained least squares method, like the second one of data set 1 and the first one of data set 2. It can be easily explained by the fact that least squares afford all interests on statistical meaning to obtain optimal solution, while constrained least squares method shares the physical meaning of the coefficients and gets a suboptimal one with a small regression-precise decrease.

TABLE II
QUALITY OF DIFFERENT FUSION METHODS ON DATA SETS

Dataset	Method	SAM (Order)	ERGAS (Order)	Q4 (Order)	Sum of Order
1	GS_AVE	1.263	1.220	0.893	
	GS_LS	1.259	1.220	0.892	
	GS_CLS	1.255	1.210	0.893	
	FIHSSA	1.562	2.260	0.835	
	PANSHARP	1.525	1.373	0.870	
2	GS_AVE	7.489 (1)	8.198 (3)	0.803 (2)	6 (2)
	GS_LS	7.627 (3)	7.955 (1)	0.803 (2)	6 (2)
	GS_CLS	7.541 (2)	8.024 (2)	0.806 (1)	5 (1)
	FIHSSA	7.745 (4)	8.684 (4)	0.802 (4)	12 (4)
	PANSHARP	10.809 (5)	10.166 (5)	0.736 (5)	15 (5)

(*): * is the quality order.

- 2) Since the lack of 850–900-nm spectral range and the corresponding relative spectral response should be considered in comparing MS with Pan of IKONOS sensors, it is wrong that the sum of coefficients is one for the average method. However, for data set 1, that of the least squares is $0.30 - 0.07 + 0.48 + 0.20 = 0.91$, and it is $0.15 + 0.27 + 0.08 + 0.37 = 0.87$ for the constrained least squares, which are sound and acceptable. The same phenomenon occurs on the other data set, which are 0.98 for least squares and 0.91 for constrained least squares.

Wald's strategy [9] is adopted to get the true reference. Original MS and Pan are downsampled to a low resolution by a factor of four, which are then fused, and quality measurements are undertaken between the fused image and the original MS. Qualitative judgments are made upon visual analysis. Quantitative indexes are spectral angle mapper (SAM), *erreur relative globale adimensionnelle de synthese* (ERGAS), global quality index (Q4) [10]. Gram–Schmidt spectral sharpening with simulated LRPs, abbreviated as GS_AVE (average), GS_LS (least squares), GS_CLS (constrained least squares) and two other homologous methods, fast intensity–hue–saturation fusion with spectral adjustment (FIHSSA) [1] and PANSHARP, provided by PCI software, are compared with each other.

For data set 1, visual comparison are shown in Fig. 5; the distinguishing difference is highlighted by the red rectangle in the bottom-left. GS_AVE and FIHSSA are much redder than the others, while PANSHARP seems to be bluer over the whole image. GS_LS and GS_CLS get the best color with the benchmark. The big gap of the fusion quality causes absolute distance on the quantitative indexes, reported in Table II. As all the three indexes are concerned, FIHSSA obtains the highest SAM, highest ERGAS, and lowest Q4, which means the worst performance, followed by PANSHARP, and then GS_LS. Although the Q4 of GS_AVE and GS_CLS are equal, SAM and ERGAS are smaller for GS_CLS than those of GS_AVE. All the evidences show the superiority of GS_CLS.

Results of data set 2 are shown in Fig. 6. There is a lot of noise on PANSHARP [the orange rectangle on Fig. 6(f)], which induces its drawback on the three indexes in Table II. The SAM, GS_AVE attains a lowest value, followed by GS_CLS, the situation among the others seems to be more complex.

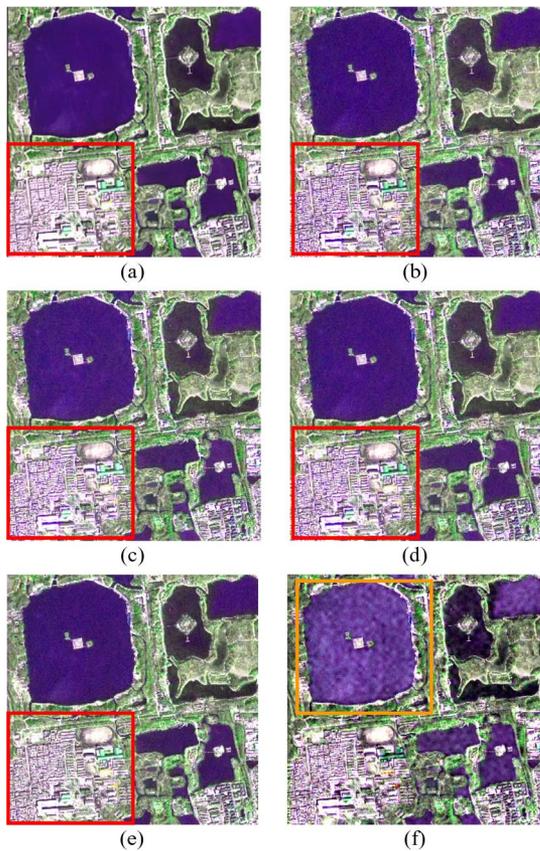


Fig. 6. Fusion experiments on data set 2. (a) Reference. (b) GS_AVE. (c) GS_LS. (d) GS_CLS. (e) FIHSSA. (f) PANSHARP.

Tiny difference could be identified in the area of buildings highlighted by the red rectangle in the bottom left, while FIHSSA is much brighter, GS_AVE is a little more purple, and the other two are almost the same as the original MS. Further comparison should be done based on quantitative scores listed in Table II. Individual comparisons are undertaken on every index, and then integrative conclusion could be achieved. In the view of ERGAS, the order is GS_LS, GS_CLS, GS_AVE, FIHSSA, and PANSHARP. As Q4 is concerned, the order is GS_CLS, GS_AVE, GS_LS, FIHSSA, and PANSHARP. The less satisfied linear regression should take the responsibility. Since near-infrared bandwidth range of 850–900 nm is excluded from MS sensors, it is the large amount of vegetation on data set 2 that reflects highly on these channel of Pan sensor brings large simulation errors.

If the order positions are ranked by one to five, with one meaning the best, a combined quality order could be measured by the sum of the order, and the method getting the lowest number performs best. As listed in the last column in Table II, GS_CLS gains the lowest sum, which indicates its highest fusion quality.

According to the analysis, to redemonstrate the superior performance for the proposal on data set 1, an additional experiment is shown briefly in Fig. 7 and quality indexes are shown in Table III. Again, GS_CLS performs best. Apart from PANSHARP and FIHSSA, since the injecting strategies of the other methods are the same, it is the regression coefficients that determine the fusion qualities.

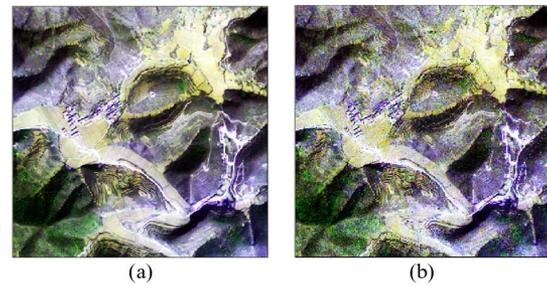


Fig. 7. Fusion result of another experiment. (a) Reference. (b) GS_CLS.

TABLE III
QUALITY OF ADDITIONAL EXPERIMENT

Method	SAM	ERGAS	Q4
GS_AVE	1.0434	0.8803	0.8809
GS_LS	1.0382	0.8749	0.8815
GS_CLS	1.0328	0.8705	0.8816
FIHSSA	1.1209	2.1477	0.8215
PANSHARP	1.4394	1.4182	0.8237

V. CONCLUSION

With an objective-oriented workflow, an adaptive and suitable threshold is found to remove high-frequency pixels from being observed in multivariate linear regression, with all low-frequency pixels being regarded equally. In addition, the fusion coefficients are obliged to be sound physically, and the fusion qualities are superior to the other four general methods. The size of subsets in SVM, that of buckets in bucketing technique, and the minimum amount of coefficient, play an important role when using the described method. Those will be investigated in the future research. In addition, images from other sensors will be tested in the future.

REFERENCES

- [1] T.-M. Tu, S. H. Ping, C.-L. Hung, and C.-P. Chang, "A fast intensity–hue–saturation fusion technique with spectral adjustment for IKONOS imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.
- [2] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pan-sharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.
- [3] W. Dou, Y. Chen, X. Li, and D. Z. Sui, "A general framework for component substitution image fusion: An implementation using the fast image fusion method," *Comput. Geosci.*, vol. 33, no. 2, pp. 219–228, Feb. 2007.
- [4] K. A. Kalpoma and J. I. Kudoh, "Image fusion processing for IKONOS 1-m color imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3075–3086, Oct. 2007.
- [5] G. A. Maria, X. Otazu, O. Fors, and A. M. Jesus, "A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1683–1691, Jun. 2006.
- [6] Y. Zhang, "Detection of urban housing development by fusing multisensor satellite data and performing spatial feature post-classification," *Int. J. Remote Sens.*, vol. 22, no. 17, pp. 3339–3355, Nov. 2001.
- [7] X. Huang, L. Zhang, and P. Li, "Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 260–264, Apr. 2007.
- [8] C. V. Tao and Y. Hu, "A comprehensive study on the rational function model for photogrammetric processing," *Photogramm. Eng. Remote Sens.*, vol. 64, no. 12, pp. 1347–1357, Dec. 2001.
- [9] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.
- [10] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.