2018

Sparse Representation over Shared Coefficients in Multispectral Pansharpening

Liuqing Chen
Department of Electronic Engineering and State Key Lab of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China.

Xiaofeng Zhang
Department of Electronic Engineering and State Key Lab of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China.

Hongbing Ma
Department of Electronic Engineering and State Key Lab of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China and College of Information Science and Engineering, Xinjiang University, Urumqi 830046, China.

Follow this and additional works at: https://tsinghuauniversitypress.researchcommons.org/tsinghua-science-and-technology

Part of the Computer Sciences Commons, and the Electrical and Computer Engineering Commons

Recommended Citation
Liuqing Chen, Xiaofeng Zhang, Hongbing Ma. Sparse Representation over Shared Coefficients in Multispectral Pansharpening. Tsinghua Science and Technology 2018, 23(03): 315-322.

This Research Article is brought to you for free and open access by Tsinghua University Press: Journals Publishing. It has been accepted for inclusion in Tsinghua Science and Technology by an authorized editor of Tsinghua University Press: Journals Publishing.
Sparse Representation over Shared Coefficients in Multispectral Pansharpening

Liuqing Chen, Xiaofeng Zhang, and Hongbing Ma

Abstract: The pansharpening process is for obtaining an enhanced image with both high spatial and high spectral resolutions by fusing a panchromatic (PAN) image and a low spatial resolution multispectral (MS) image. Sparse Principal Component Analysis (SPCA) method has been proposed as a pansharpening method, which utilizes sparse coefficients and over-complete dictionaries to represent the remote sensing data. However, this method still has some drawbacks, such as the existence of the block effect. In this paper, based on SPCA, we propose the Sparse over Shared Coefficients (SSC), in which patches are extracted with a sliding distance of 1 pixel from a PAN image, and the MS image shares the sparse representation coefficients trained from the PAN image independently. The fused high-resolution MS image is reconstructed by K-SVD algorithm and iterations, and residual compensation is applied when the down-sampling constraint is not satisfied. The simulated experiment results demonstrate that the proposed SSC method outperforms SPCA and improves the overall effectiveness.

Key words: pansharpening; sparse representation; shared coefficients; iteration

1 Introduction

In remote sensing systems, the limitations of data transmission greatly affect spatial and spectral resolutions; therefore, optical systems provide either an image with a high spatial resolution but a low spectral resolution, such as panchromatic (PAN) images, or an image with a high spectral resolution but a low spatial resolution, such as multispectral (MS) images with several bands and hyperspectral (HS) images with hundreds of bands[1]. However, images with high spatial and spectral resolutions are required in many fields, such as agricultural production, environmental protection, and land utilization. To overcome these limitations and obtain the required image, the spatial and spectral information of the PAN and MS images have to be integrated, a process called pansharpening[2]. This would enhance the spatial resolution of MS data and provide an enhanced image with high spatial and spectral resolutions[3].

In recent decades, pansharpening has gained much attention, as shown by the existence of diverse literature. Pansharpening algorithms have been surveyed in detail in Ref. [4], most of which can be divided into three types: spatial fusion, spectral substitution, and optimization reconstruction methods.

In the spatial fusion methods, spatial details are extracted from the PAN image and are then injected into different MS bands, which contain algorithms such as Pradines, Price, and Brovery. In Pradines, the pixels in a patch of the PAN image are regarded as the weights of the corresponding MS pixels[5]. Compared to Pradines, Price applies linear fittings to reconstruct a middle image.
whose pixels are used as the weights of the corresponding MS image\cite{6}. In contrast, for Brovey, the registered pixels in various MS bands are regarded as the weights and are then injected into the corresponding patch of the PAN image\cite{7}. Smoothing Filter-based Intensity Modulation (SFIM)\cite{8} and Laplacian pyramid\cite{9} both belong to this class.

Next, spectral substitution methods try to split the spatial information and the spectral information, and then substitute the spatial information for the PAN image, which contains Intensity-Hue-Saturation (IHS)\cite{10}, Principal Component Analysis (PCA)\cite{11}, Wavelet Transform (WT)\cite{12}, and Gram-Schmidt (GS)\cite{13} spectral sharpening. The Band-Dependent Spatial Detail (BDSD)\cite{14} algorithm falls within the class and is one of the most advanced schemes.

Recently, optimization reconstruction has drawn lots of attention, in which reconstruction based on sparse representation is widely utilized in image processing. There are two viewpoints for sparse representation in remote sensing images. Li et al.\cite{15} tried to use over-complete dictionaries to model PAN and MS images, wherein different MS bands share the sparse representation coefficients, a method called Sparse over Learned Dictionaries (SLD). On the contrary, Zhu and Bamler\cite{16} proposed Sparse Fusion of Images (SFI), where the PAN and MS images under the same spatial resolution share the over-complete dictionaries.

Moreover, Sparse Principal Component Analysis (SPCA) method has been introduced in our previous research, wherein shared sparse representations and PCA are combined\cite{17}. There are two disadvantages of SPCA: the block effect and the limitation of PCA when they are insufficient bands.

In this paper, we propose the SSC based on SPCA. As the core concept in SSC, we prove experimentally that if we learn the dictionary and sparse representation coefficients from the PAN image independently, which are shared in the MS image to reconstruct its dictionary, we would get an estimated target MS image closer to the actual target. Then the pansharpening process based on sparse representation is transformed to an optimization problem, where the cost function is the reconstruction error, so that it can be solved by iterations. The SSC method outperforms the SPCA method in almost all respects. In SSC, we extract patches from PAN image with a sliding distance of 1 pixel, which produces a refined sampling and avoids the block effect.

The rest of this paper is organized as follows. In Section 2, we briefly review the SPCA method. The sparsity and the proposed SSC method are described in Section 3. Section 4 presents the results and discussion. Finally, we state our conclusions in Section 5.

2 SPCA

First of all, we introduce the notations used in this paper. We represent a remote sensing image as a matrix, where each column represents a band of an image (two-dimensional image turned into a column vector), and different columns represent different bands. Therefore, we use the following matrices in this paper:

- The matrices $P \in \mathbb{R}^{n \times 1}$, $Y_M \in \mathbb{R}^{m \times m \lambda}$ represent the observed PAN and MS images, respectively. The observed MS image refers to the low spatial resolution MS image (LR MS image). The total number of pixels in the PAN image is denoted by $n$, and the total number of pixels of one band of the MS image is denoted by $m$, and $m \lambda$ refers to the number of bands of the $Y_M$ image. Generally $m < n$.
- Also, $X_M \in \mathbb{R}^{n \times m \lambda}$ represents the target image, which includes the high spatial resolution MS image (HR MS image) with $m \lambda$ bands and $n$ pixels for one band, and $\tilde{X}$ represents the estimated image for $X_M$.
- We denote $x_{kM}$ as the $k$-th column of $X_M$, $y_{kM}$ as the $k$-th column of $Y_M$, and $x_P$ as the column of $P$.

The SPCA method uses sparse representation in processing of remote sensing image data. According to sparse representation theory, the column vector $x \in \mathbb{R}^n$ can be expressed as a linear combination of $N$ $n$-dimensional vectors ($n < N$), which is formulated by

$$x = D\alpha$$

where columns for the dictionary $D \in \mathbb{R}^{n \times N}$ are $n$-dimensional vectors, and $\alpha \in \mathbb{R}^n$ contains the coefficients for the dictionary.

When applying sparse representation to PAN and MS images, there are two assumptions: (1) the MS and PAN images share the same sparse coefficients $\alpha$, and have different dictionaries; (2) the LR MS image is the downsampling version of the HR MS image. Thus, we have

$$x_P = D_P\alpha + n_P$$

$$x_{kM}^M = D_{kM}\alpha + n_{kM}$$
where $D_P$ and $D_M^*$ are the dictionaries of the PAN image and the $k$-th band of the HR MS image, respectively, and $n_P$ and $n_M^*$ refer to noises.

Denoting $S \in \mathbb{R}^{3 \times n}$ as the down-sampling matrix from the HR MS image to the LR MS image, where $\gamma = \sqrt{n/m}$ is the down-sampling factor, we have

$$y_M^k = Sx_M^k = SD_M^*\alpha + Sn_M^k = D_M^{*\gamma}\alpha + n_M^{k\gamma} \quad (4)$$

where $D_M^{*\gamma}$ is the dictionary for the $k$-th band of the LR MS image and $n_M^{k\gamma}$ is the noise.

Generally, the LR MS image is assumed as the degraded version of the HR MS image, through MTF-filter or down-sampling. Here, for convenience, we assume that the LR MS image is the down-sampling version of the HR MS image. We aim to find the MTF-filter; therefore, we do not need to calculate the exact matrix, as the MTF-filter is hidden in the pansharpening process.

Furthermore, the LR MS images share the same sparse coefficients $\alpha$ with the PAN image and the HR MS image. Let $y_M = \left[ [y_M^1]^T, [y_M^2]^T, \ldots, [y_M^n]^T \right]^T$ and $D_M = \left[ [D_M^1]^T, [D_M^2]^T, \ldots, [D_M^n]^T \right]^T$, the unified form is given by

$$y_M = D_M^*\alpha + n_M^\gamma \quad (5)$$

Therefore, we get the optimization problem as follows:

$$\begin{align*}
\min_{\alpha} \| \alpha \|_0,
\text{s.t.,} \quad \| (\lambda x_P + \lambda D_M) - (D_M^*\alpha) \|_2^2 \leq \epsilon^2
\end{align*} \quad (6)$$

where $\lambda$ represents the relative tolerance of residual in PAN image compared to LR MS image, and $\epsilon \geq 0$ is the error tolerance.

The pansharpening optimization problem is divided into the following three steps in SPCA\cite{17};

- Based on the joint training of $P$ and $Y_M$, we obtain the over-complete dictionaries $D_P$ and $D_M^*$ and the sparse representation coefficients $\alpha$.
- PCA is used to construct the dictionary $D_M$ of $X_M$ from $D_P$ and $D_M^*$, and it is assumed that the PCs for $Y_M$ is the down-sampling version of the PCs for $X_M$.
- Residual compensation is applied after constructing $X_M^*$ with the sparse model.

### 3 SSC

Although the SPCA method is acceptably stable and effective, it has two disadvantages: First, when we train the dictionaries, the image patches are extracted from LR MS image with a sliding distance of 1 pixel, while the PAN image patches are extracted with sliding distance $\gamma$. The sampling of the PAN image is not refined, which may cause information loss, and hence, the block effect. Second, PCA reduces the generality of the PAN image and the MS bands in texture and spatial information, which is the innovation point of SPCA method. However, if the number of bands is small, the similarity and distinction between MS bands and PAN image are not obvious, and the weights of different bands would not be estimated accurately. Therefore, PCA becomes ineffective, and then SPCA method does not perform ideally. In this paper, we do not consider PCA, and propose a theory with fewer hypotheses and broader applications. In the rest of this section, we introduce the SSC method, which outperforms the SPCA.

First, we describe the sparsity, which usually denotes non-zero entries of a sparse vector or signal. However, the definition of sparsity for an image is not clear. In this paper, we relate the sparsity of an image to the probability of reconstructing the image by sparse representation with little error, which differs from the general definition. On this occasion, when an image is reconstructed by sparse representation, a small error between the original and reconstructed images means that the original image is sparse.

If we confirm that the HR MS image is more sparse, we could input it as a variable into the original optimization problem, thereby making the target image more and more sparse through continuous optimizations; this is the original consideration to improve SPCA.

In the experiments of Section 5, we test two patterns: training PAN individually and jointly. We found that if we learn the dictionary and the sparse representation coefficients from the PAN image independently, which are shared in the MS image, and we then reconstruct the dictionary, the reconstructed HR MS image would be closer to the actual target image.

Therefore, we model the pansharpening based on sparse representation as an optimization problem with the cost function as the reconstruction error, so that it can be solved by iterations. The optimization problem for $k$-th band of the HR MS image is

$$\begin{align*}
\min_{D_M^k, x_M^k} \sum_{b} \| R_b x_M^k - D_M^k \alpha_b \|_2^2, \\
\text{s.t.,} \quad Sx_M^k = y_M^k
\end{align*} \quad (7)$$

where $b$ refers to the serial number of the sampling block.
Thus $R_b$ represents sampling matrix of the $b$-th block, and $\alpha_b$ represents the sparse representation coefficients, which are trained from PAN image.

This is a convex optimization problem, which can be solved by Lagrange operators. When the size of the images is too large, considering the computational burden, we propose an iterated algorithm named SSC as follows:

1. First, we extract patches with a sliding distance of 1 pixel from a PAN image, and then apply the K-SVD algorithm[18] to obtain its dictionary and the shared sparse representation coefficients $\alpha_b$.

2. For each band, we initialize $x_M = S_r y_M$, where $S_r$ is the nearest-neighbor up-sampling matrix. Then, we repeat the following steps until convergence.

3. Input $x_M$, and optimize

$$\min_{D_M} \sum_b \| R_b x_M - D_M \alpha_b \|_2^2$$

to solve the dictionary $D_M$. Reconstruct $x_M$ by $D_M \alpha_b$, and for one pixel, if it is sampled repeatedly, take the average of the results.

4. If the down-sampling constraint for $x_M$ is not satisfied (i.e., the LR MS image is the down-sampling version of the HR MS image), perform $x_{M+1} = x_M + S_r (y_M - S x_M)$.

5. Let $x_M = x_M^{k+}$ and $D_M = D_M^{k+}$. If convergence occurs, stop; otherwise jump to step (3).

SSC extracts patches from a PAN image with a sliding distance of 1 pixel, which produces a refined sampling and increases the number of training samples. In theory, the size of the image patches can be adjusted for SSC, whereas for SPCA, it is a multiple of the down-sampling factor, which consequently causes the block effect.

There is generality in the texture and spatial information for the PAN image and MS bands, which is not evident in less spectral bands. SSC depicts this generality as here, the MS images share the sparse representation coefficients trained from an PAN image rather than PCA.

4 Experimental Results and Discussion

4.1 Measurements

Before explaining the results, we introduce the used quality measures: Correlation Coefficients (CC), Root-Mean-Square Error (RMSE), Multispectral Entropy (ME), and Average Gradient (AG). The first two need the reference images to be evaluated. As a spectral measurement, ME reflects the entropy of an MS image with $m$ bands, where 1 pixel value could be represented by $\beta$ bits. According to Shannon information theory, ME is formulated as

$$ME = - \sum_{i_1, \ldots, i_m=0}^{2^\beta-1} P_{i_1,\ldots, i_m} \log_2 P_{i_1,\ldots, i_m} \quad (8)$$

where $i_k$ denotes the pixel value of the $k$-th band of the MS image, and $P_{i_1,\ldots, i_m}$ refers to the corresponding probability with the pixel value $(i_1,\ldots, i_m)$. Specifically, images with abundant spectral information have large ME indexes.

Moreover, AG is proposed for spatial measurement, which is represented as

$$AG = \frac{1}{mn} \sum_{i=1}^{mn} \left( \frac{\Delta P_i(x,y)}{\Delta x} \right)^2 + \left( \frac{\Delta P_i(x,y)}{\Delta y} \right)^2 \quad (9)$$

This index measures the texture and detail information, whereby a higher AG indicates more image details.

4.2 Datasets

Landsat 8 and QuickBird datasets were used in the experiment. The sizes of the images are represented in Table 1.

For simulated experiments, the original PAN and MS images are degraded at the same rate. As a result, we regard the original images as the reference and the degraded images as the inputs of pansharpening algorithms.

4.3 Pilot experiments

In this subsection, we use three images: PAN image, original MS image (the same spatial resolution as PAN image), and blurred MS image (achieved by downsampling and nearest neighbor interpolation of the original MS image).

In Table 2, the performance is evaluated on Landsat8’s PAN image and B2 band image. The distance between the original and blurred MS images is 191.55.

| Dataset | Ori. PAN image | Ori. MS image | Degraded PAN image | Degraded MS image | Down-sampling factor $\gamma$ | Number of the used bands | Redundancy rate of dictionaries |
|---------|----------------|---------------|-------------------|-------------------|--------------------------|-------------------------|-----------------------------|
| Landsat 8 | 1024 $\times$ 1024 | 512 $\times$ 512 | 512 $\times$ 512 | 256 $\times$ 256 | 2 | 6 | 4 |
| QuickBird | 2048 $\times$ 2048 | 512 $\times$ 512 | 512 $\times$ 512 | 128 $\times$ 128 | 4 | 4 | 4 |
In one case, we learned the dictionary and the sparse representation coefficients from the PAN image independently, which are then shared by the original and blurred MS images to train their dictionary. In the table, the reconstructed RMSE of the original MS image is 42.01, and 59.49 for the blurred MS image. Considering the definition of sparsity in Section 3, the smaller the error of the reconstructed image, the more sparse original image becomes. Hence, the original MS image is more sparse than the blurred one. When we reconstructed the blurred MS image by its dictionary and sparse coefficients, the distance to the original MS image reduced from 191.55 to 77.21. In another case, we joined trained dictionaries of PAN and the blurred images and found that though reconstructed RMSE of the blurred image is smaller, the distance to the original MS image is 119.60, greater than 77.21.

For pansharpening, if we treat the LR MS image as the blurred image in the optimization problem, the LR MS image (the blurred image) could approach the HR MS image (the original MS image) by sharing PAN sparse representation coefficients.

The ultimate goal of the optimization is to make the MS image more and more sparse in the iteration process with the down-sampling constraint satisfied. In most cases, blurred images are more sparse than clear images. Therefore, we need to take the sparse representation coefficients of the PAN image as the constraint of the HR MS image. Then the fuzzy MS image becomes relatively more sparse and the optimization process makes it iteratively approximate to clear the MS image.

### 4.4 Simulated experiments

The numerical results achieved by PCA, SFIM, Price, BDSD, SFI, SLD, SPCA, and SSC methods on the two datasets are reported in Tables 3 and 4. In the experiments, the number of iterations is between 30 and 60, which is acceptable for the time.

In Table 3, SSC performs best for almost all indexes, including CC, ERGAS, and SAM, especially for CC, where it is 0.7% better than SPCA. This is a great progress, which reveals the reasonability and practicability of the

| Table 2: Sparsity of PAN and MS images. |
|----------------------------------------|
| Reconstructed RMSE | PAN | MS | Blurred MS |
|---------------------|-----|----|------------|
| RMSE                | Dist. to MS |
| Shared coefficients | 28.86 | 42.01 | 59.49 | 77.21 |
| Joint PAN and blurred | 39.41 | – | 16.58 | 119.60 |

| Table 3: Evaluation on simulated LandSat8. |
|----------------------------------------|
| Index | Ref. image | PCA | SFIM | Price | BDSD | SFI | SLD | SPCA | SSC |
|-------|------------|-----|------|-------|------|-----|-----|------|-----|
| B2    | 1          | 0.939 | 0.953 | 0.977 | 0.974 | 0.967 | 0.974 | 0.984 | 0.989 |
| B3    | 1          | 0.932 | 0.949 | 0.978 | 0.969 | 0.959 | 0.979 | 0.979 | 0.983 |
| B4    | 1          | 0.927 | 0.949 | 0.978 | 0.969 | 0.959 | 0.975 | 0.975 | 0.983 |
| B5    | 1          | 0.953 | 0.902 | 0.970 | 0.997 | 0.979 | 0.948 | 0.948 | 0.970 |
| B6    | 1          | 0.937 | 0.942 | 0.975 | 0.986 | 0.973 | 0.958 | 0.974 | 0.979 |
| B7    | 1          | 0.916 | 0.948 | 0.975 | 0.967 | 0.961 | 0.961 | 0.978 | 0.986 |
| Avg   | 1          | 0.934 | 0.940 | 0.976 | 0.976 | 0.976 | 0.961 | 0.977 | 0.984 |
| B2    | 0          | 16.71 | 14.70 | 9.86 | 10.97 | 11.69 | 10.40 | 8.04 | 6.63 |
| B3    | 0          | 17.35 | 15.20 | 9.72 | 12.07 | 13.01 | 9.47 | 8.53 | 6.71 |
| B4    | 0          | 17.88 | 15.28 | 9.74 | 12.26 | 13.10 | 10.41 | 8.45 | 6.59 |
| B5    | 0          | 14.41 | 20.89 | 11.43 | 4.30 | 9.55 | 14.98 | 12.81 | 11.40 |
| B6    | 0          | 17.52 | 16.30 | 10.27 | 8.22 | 10.71 | 13.35 | 10.58 | 9.44 |
| B7    | 0          | 19.27 | 16.03 | 10.09 | 11.19 | 12.13 | 17.40 | 9.99 | 8.09 |
| Avg   | 0          | 17.19 | 16.40 | 10.19 | 9.84 | 11.70 | 12.67 | 9.73 | 8.14 |
| B2    | 14.07 | 15.84 | 19.14 | 17.12 | 17.64 | 10.41 | 11.11 | 13.42 | 14.00 |
| B3    | 16.57 | 16.88 | 20.11 | 18.24 | 20.31 | 11.95 | 13.54 | 15.15 | 15.89 |
| B4    | 16.54 | 17.43 | 20.35 | 18.42 | 20.24 | 11.96 | 17.33 | 15.20 | 15.99 |
| B5    | 13.04 | 9.54 | 18.63 | 9.13 | 14.62 | 9.54 | 14.98 | 12.18 | 9.26 |
| B6    | 14.51 | 14.78 | 20.19 | 13.75 | 17.50 | 10.49 | 8.65 | 13.33 | 12.93 |
| B7    | 16.46 | 17.46 | 21.26 | 17.30 | 20.38 | 12.03 | 11.23 | 15.63 | 15.37 |
| RMSE  | 0        | 1.70 | 4.82 | 2.31 | 3.35 | 4.16 | 4.85 | 1.09 | 1.77 |
| ERGAS | 0        | 6.783 | 6.492 | 4.009 | 4.020 | 4.619 | 5.096 | 3.853 | 3.236 |
| SAM   | 0        | 3.529 | 3.002 | 2.341 | 1.649 | 2.454 | 3.660 | 2.517 | 2.126 |
Table 4  Evaluation on simulated QuickBird.

| Index | Ref | PCA  | SFIM | Price | BDSD | SFI  | SLD  | SPCA | SSC |
|-------|-----|------|------|-------|------|------|------|------|-----|
| CC    | B   | 1    | 0.948| 0.940 | 0.972| **0.991** | 0.948 | 0.940 | 0.978| 0.981|
|       | G   | 1    | 0.958| 0.943 | 0.978| **0.989** | 0.937 | 0.985 | 0.984| 0.987|
|       | R   | 1    | 0.962| 0.944 | 0.981| **0.989** | 0.934 | 0.987 | 0.985| **0.989**|
|       | NI  | 1    | 0.958| 0.926 | 0.982| **0.988** | 0.900 | 0.979 | 0.980| 0.984|
|       | Avg | 1    | 0.956| 0.938 | 0.978| **0.989** | 0.929 | 0.972 | 0.982| 0.985|
| RMSE  | B   | 0    | 13.09| 14.29 | 9.64 | **5.76**  | 12.90 | 13.80 | 8.41 | 7.89 |
|       | G   | 0    | 11.41| 13.48 | 8.41 | **5.84**  | 13.01 | 6.91  | 7.02 | 6.33 |
|       | R   | 0    | 11.34| 14.04 | 8.17 | 6.40      | 14.88 | 6.64  | 7.00 | **6.14**|
|       | NI  | 0    | 12.46| 16.44 | 8.21 | **6.78**  | 18.88 | 8.99  | 8.57 | 7.62 |
|       | Avg | 0    | 12.08| 14.56 | 8.61 | **6.19**  | 15.10 | 9.09  | 7.77 | 7.00 |
| AG    | B   | 7.74 | 9.73 | 10.90 | 10.01| 9.79      | 6.61  | 3.54  | 7.31 | **7.33**|
|       | G   | 7.98 | 10.01| 10.69 | 10.83| 9.59      | 7.06  | 7.04  | **8.24**| 8.28 |
|       | R   | 8.71 | 10.42| 11.31 | 11.55| 10.75     | 7.73  | 9.12  | 9.05 | **8.38**|
|       | NI  | 12.84| 10.28| 11.58 | 11.83| 14.79     | 9.06  | **12.49**| 11.30| 11.25 |
|       | RMSE| 0    | 2.09 | 2.53  | 2.37 | 1.92      | 2.08  | 2.17  | 0.84 | 0.85 |
| ERGAS | 0   | 2.411| 2.914| 1.721 | 1.390| 3.030     | 1.903 | 1.549 | 1.411|
| SAM   | 0   | 2.822| 2.778| 2.251 | 2.521| **0.827** | 3.153 | 2.621 | 2.217| 2.094|

Theoretical assumptions. In terms of the AG index, SSC and SPCA perform equally. The AG index for Price and BDSD is very high in some bands, such as B6 and B7. The higher the AG index, the more abundant the texture information is; however, Price and BDSD is over-fused for almost all bands and unstable in terms of the AG index. On the contrary, though the AG index of SSC is not the highest, it is the most close to the reference MS image, especially for bands B2, B3, and B4. For this dataset, SSC outperforms BDSD and is much better than other existing SR-based methods.

For QuickBird dataset in Table 4, the CC index outperforms SPCA by 0.3%, while the ERGAS and SAM indexes are both enhanced. In terms of the AG index, SSC method is slightly over-fused. We notice that for this dataset, BDSD is slightly better than our method, except in the AG index.

Visual analysis of SSC method compared with PCA, SFIM, Price, BDSD, SFI, SLD, and SPCA can be performed via Figs. 1 and 2. SSC method provides good spatial results with little spectral distortions.

We enlarge the fused results with SPCA and SSC methods, as presented in Figs. 3 and 4. There is obviously a seam line for the SPCA result, while it is smoother for SSC, which means that the block effect in SSC is weakened.

![Fig. 1 Simulated images and fused results on B4, B3, and B2 with different methods on LandSat8. (a) Original MS image, (b) degraded PAN image, (c) degraded MS image, (d) PCA method, (e) SFIM method, (f) Price method, (g) BDSD method, (h) SFI method, (i) SLD method, (j) SPCA method, (k) SSC method.](image-url)
The advantages of the SSC method over the SPCA method include: (1) The sampling distance on an HR image is 1, which refines the image block, increases the number of training samples, and theoretically makes the size of the image block adjust arbitrarily. However, in SPCA, the size of the HR image block must be a multiple of the down-sampling rate, which makes the block effect unavoidable. (2) The SSC method utilizes the shared sparse representation coefficients from the PAN image deeper, making our theory more interesting. Actually the dictionary of each band in the remote sensing image is related to each other in texture; however this may not be obvious for a small number of bands. The SSC method hides the inter-dictionary association in the down-sampling constraint condition, and the correlation of the images is described by sharing the sparse representation coefficient from the PAN image. These concise ideas depict the similarity of image blocks in the same band of remote sensing images and the correlation and similarity between different bands. Constraints of common sparse representation coefficients and over-complete dictionaries make the theoretical framework more in line with the actual remote sensing image.

5 Conclusion

In this paper, a pansharpening method, SSC, is proposed. As the core concept in SSC, we show by experiment that if we learn the dictionary and the sparse representation coefficients from the PAN image independently, which are then shared by the MS image to reconstruct its dictionary, we would get an estimated target MS image that is closer to the actual target MS image. The pansharpening process based on sparse representation is then transformed into an optimization problem with the cost function being the reconstruction error, so that it can be solved by iterations. SSC outperforms SPCA in almost all respects, such as CC and RMSE indexes. In SSC, we extract patches with a sliding distance of 1 pixel from PAN image, which makes a refined sampling and avoids the block effect.

References

[1] Y. Fan, D. Ni, and H. Ma, Hyperdb: A hyperspectral land class database designed for an image processing system, *Tsinghua Science and Technology*, vol. 22, no. 1, pp. 112–118, 2017.
[2] L. Wald, Some terms of reference in data fusion, *IEEE Transactions on Geoscience & Remote Sensing*, vol. 37, no. 3, pp. 1190–1193, 2009.

[3] C. Pohl and J. L. V. Genderen, Review article multisensor image fusion in remote sensing: Concepts, methods and applications, *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 823–854, 1998.

[4] G. Vivone, L. Alparone, J. Chanussot, M. D. Mura, A. Garzelli, and G. Licciardi, A critical comparison of pansharpening algorithms, in *Proc. IEEE Conf. Geoscience and Remote Sensing Symposium*, Quebec City, Canada, 2014.

[5] D. Pradines, Improving spot images size and multispectral resolution, in *Proc. SPIE 0660, Earth Remote Sensing using the Landsat Thematic Mapper and SPOT systems, Proceeding SPIE Conference*, Innsbruck, Austria, 1986.

[6] J. C. Price, Combining panchromatic and multispectral imagery from dual resolution satellite instruments, *Remote Sensing of Environment*, vol. 21, no. 2, pp. 119–128, 1987.

[7] V. Vijayaraj, N. H. Younan, and C. G. O‘Hara, Quantitative analysis of pansharpened images, *Optical Engineering*, vol. 45, no. 4, pp. 828–828, 2006.

[8] J. G. Liu, Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details, *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3461–3472, 2000.

[9] P. J. Burt and E. H. Adelson, The Laplacian pyramid as a compact image code, *Readings in Computer Vision*, vol. 31, no. 4, pp. 671–679, 1987.

[10] V. K. Shettigara, A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set, *Photogrammetric Engineering & Remote Sensing*, vol. 58, no. 5, pp. 561–567, 1992.

[11] C. Psjr, S. C. Sides, and J. A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, *Photogrammetric Engineering & Remote Sensing*, vol. 57, no. 3, pp. 265–303, 1991.

[12] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods, *IEEE Transactions on Geoscience & Remote Sensing*, vol. 43, no. 10, pp. 2376–2385, 2005.

[13] C. A. Laben and B. V. Brower, Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, US Patent US6011875, 2000.

[14] A. Garzelli, F. Nencini, and L. Capobianco, Optimal nmse pan sharpening of very high resolution multispectral images, *IEEE Transactions on Geoscience & Remote Sensing*, vol. 46, no. 1, pp. 228–236, 2007.

[15] S. Li, H. Yin, and L. Fang, Remote sensing image fusion via sparse representations over learned dictionaries, *IEEE Transactions on Geoscience & Remote Sensing*, vol. 51, no. 9, pp. 4779–4789, 2013.

[16] X. X. Zhu and R. Bamler, A sparse image fusion algorithm with application to pan-sharpening, *IEEE Transactions on Geoscience & Remote Sensing*, vol. 51, no. 5, pp. 2827–2836, 2013.

[17] X. Zhang, D. Ni, Z. Gou, and H. Ma, Sparse representation and PCA method for image fusion in remote sensing, in *Proc. IEEE Conf. International Conference on Control, Automation and Robotics*, Hong Kong, China, 2016.

[18] M. Aharon, M. Elad, and A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, *IEEE Trans on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

**Liuqing Chen** received the BEng degree from Tsinghua University, China, in 2015. She is currently a master student at the Department of Electronic Engineering, Tsinghua University, China. Her research interests include remote sensing, image processing, and pattern recognition.

**Hongbing Ma** received the PhD degree from Peking University, China in 1999. He is currently an associate professor with the Department of Electronic Engineering, Tsinghua University, China. His research interests include image processing, pattern recognition, and spatial information processing and application.

**Xiaofeng Zhang** received the BEng and BS degree in 2013 and 2016, respectively, from Tsinghua University in China. His research interests include image fusion and remote sensing.