

## Dictionary Learning Based Pan-Sharpening

Liu, D.; Boufounos, P.R.

TR2012-013 March 2012

### Abstract

Pan-sharpening is an image fusion process in which high resolution (HR) panchromatic (Pan) imagery is used to sharpen the corresponding low resolution (LR) multi-spectral (MS) imagery. Pan-sharpened MS images generally have high spatial resolutions, but exhibit color distortions. In this paper, we propose a dictionary learning based pan-sharpening process to reduce the color distortion caused by the interpolation of the MS imagery. Instead of interpolating the LR MS image before fusion, we generate an improved MS image which is sparse with respect to a dictionary learned from the image data. Our experiments on degraded QuickBird and IKONOS images demonstrate that the distortion in the MS images produced using our approach is significantly reduced.

*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.



# DICTIONARY LEARNING BASED PAN-SHARPENING

Dehong Liu, Petros T. Boufounos

Mitsubishi Electric Research Labs, 201 Broadway, Cambridge, MA 02139  
{liudh,petrosb}@merl.com

## ABSTRACT

Pan-sharpening is an image fusion process in which high resolution (HR) panchromatic (Pan) imagery is used to sharpen the corresponding low resolution (LR) multi-spectral (MS) imagery. Pan-sharpened MS images generally have high spatial resolutions, but exhibit color distortions. In this paper, we propose a dictionary learning based pan-sharpening process to reduce the color distortion caused by the interpolation of the MS imagery. Instead of interpolating the LR MS image before fusion, we generate an improved MS image which is sparse with respect to a dictionary learned from the image data. Our experiments on degraded QuickBird and IKONOS images demonstrate that the distortion in the MS images produced using our approach is significantly reduced.

**Index Terms**— Pan-sharpening, Dictionary learning, K-SVD, Sparse representation.

## 1. INTRODUCTION

A large number of satellite imaging systems, such as QuickBird, IKONOS and ALOS, produce two kinds of imagery: panchromatic (Pan) and multi-spectral (MS). Pan imagery provides high spatial resolution but no color information, whereas MS imagery provides color spectrum information but reduced spatial resolution. For a variety image applications, such as target and material identification, it is necessary to obtain MS images with both high spatial and spectral quality. To generate those images, pan-sharpening techniques have been developed to fuse MS imagery and Pan imagery and produce high spatial resolution MS imagery.

A large number of pan-sharpening methods have been developed in the past decades, such as Intensity-Hue-Saturation (IHS), Principal Component Analysis (PCA), wavelet based fusion, variational methods and others [1–3]. Although these approaches seem quite different, their basic steps are very similar. First, the low resolution (LR) MS image is interpolated into the same size as the high-resolution (HR) Pan image. Second, the high frequency component of the Pan image is injected in the interpolated MS image in a method-dependent way. For instance, in the IHS-based method, the intensity (I) component of the interpolated MS image is replaced by the Pan image, whereas in the PCA-based method, the principal component of the interpolated MS image is replaced by the Pan image. The last step is to transform the fused data back to the image domain to achieve a fused MS image.

Pan-sharpened MS images generally have good spatial resolutions, but often exhibit color distortions [4]. To improve color accuracy and reduce color distortion, a variety of algorithms have been developed, each specific to a particular fusion technique or image set [2]. Most of these algorithms focus on how to inject the high frequency component of the Pan image in the interpolated MS image,

paying little attention on the distortion caused by interpolating the MS at the initialization of the fusion process.

Recent developments in compressive sensing, sparse representations and dictionary learning have provided us with new tools to address this problem. Specifically, dictionary learning has already been applied to pan-sharpening problems [5, 6], where HR Pan and LR MS images are treated as compressive measurements. Using sparsity in a learned dictionary as a regularization prior, the HR MS images can be recovered by solving a sparse recovery problem. However, these methods require a large amount of HR MS images to train the dictionary, which are typically not available in practice.

In this paper, we propose a dictionary learning (DL) based pan-sharpening process aiming to reduce the color distortion caused by interpolating the MS image, without requiring a large training set. Compared to previous pan-sharpening methods, the novelty of our DL based pan-sharpening lies in three aspects: (1) it focuses on generating a MS image with improved resolution using DL instead of simply interpolating the LR MS image. Thus, it can be combined with any existing pan-sharpening algorithm that is initialized by interpolating a MS image; (2) it does not require a large amount of HR MS images to learn dictionary, only the images to be sharpened; (3) it provides an iterative algorithm to update the MS image according to image sparsity and consistency, yielding a MS image with less distortion. In addition, our DL based process can be applied to hyperspectral data with minimal modifications.

The next section provides some background on DL, as well as details on our DL-based pan-sharpening method. In Section 3, we provide experimental validation on degraded QuickBird and IKONOS images with quantitative analysis. Section 4 discusses the results and concludes.

## 2. DICTIONARY LEARNING BASED PAN-SHARPENING

### 2.1. Background

Dictionary learning is the process of learning a set of vectors (atoms) which can be used to sparsely represent a signal. Consider a signal  $\mathbf{x} \in \mathbf{R}^{n \times 1}$  represented by  $\mathbf{x} = \mathbf{D}\alpha$ , where  $\mathbf{D} \in \mathbf{R}^{n \times k}$  is a dictionary that contains the atoms as its columns, and  $\alpha \in \mathbf{R}^{k \times 1}$  is the representation vector. The structure exhibited in large variety of natural signals of interest can be captured by assuming that  $\alpha$  is a sparse vector under appropriate choice of the dictionary  $\mathbf{D}$ , i.e., a linear combination of a few atoms from  $\mathbf{D}$  are able to represent the signal  $\mathbf{x}$  with small error. For example, natural images can be sparsely represented in carefully designed dictionaries, such as the DCT or a wavelet basis. Designed dictionaries typically offer the advantage of efficient implementation and signal-independent design.

An alternative choice is to learn the dictionary from a set of source signals (training data)  $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_L\}$ . The basic idea of dictionary learning is to learn a dictionary that produces the best

sparse representations for the training data. If subsequent observed signals  $\{\mathbf{x}\}$  (field acquired or testing data) exhibit similar structure, they will also produce a good sparse representation in the learned dictionary. Dictionary learning has demonstrated significant improvements in image de-noising, image inpainting, and other applications [7, 8].

To learn the dictionary from the training data, typical algorithms such as the K-SVD [7] attempt to solve the following problem:

$$\tilde{\mathbf{D}} = \operatorname{argmin}_{\mathbf{D}, \alpha} \|\mathbf{Y} - \mathbf{D}\alpha\|_F^2 \text{ s.t. } \forall l \|\alpha_l\|_0 \leq T_0, \quad (1)$$

where  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_L]$  is the set of sparse coefficient vectors corresponding to the  $L$  training vectors,  $T_0$  is the desired sparsity level, and  $\|\cdot\|_F$  denotes the Frobenius norm.

## 2.2. Learning and using the dictionary

The objective of pan-sharpening is to reconstruct a HR MS image given a HR Pan image and a LR MS image. Exploiting the fact that different materials exhibit different colors and Pan values, we explore the structure of color spectrum using the principles of dictionary learning. To achieve this goal, we consider signal vectors coupling color and Pan features extracted from the given MS and Pan images  $\mathbf{x} = [(\mathbf{x}^{MS})^T (\mathbf{x}^{Pan})^T]^T$ . Since the given MS image has low resolution, some of the high-resolution image coefficients include only Pan information  $\mathbf{x}^{Pan}$  but no color information  $\mathbf{x}^{MS}$ . Using the signal vectors with no missing data—denoted using  $\mathbf{Y} \subset \{\mathbf{x}\}$ —we learn a dictionary to represent the signal structure. We use the dictionary to fill in the missing values of  $\mathbf{x}^{MS}$  by treating the corresponding Pan information  $\mathbf{x}^{Pan}$  as testing data to be sparsely represented in the learned dictionary.

We consider a pan-sharpening problem where one LR MS image and one HR Pan image are available and are well registered. Without loss of generality, we assume the LR MS image is composed of  $M$  bands, each of size  $I \times J$ , and the Pan image is of size  $NI \times NJ$ , having  $N$  times the resolution of the LR MS image (for QuickBird and IKONOS satellite imaging systems,  $M = 4$  and  $N = 4$ ). Therefore, each pixel in the LR MS image corresponds to an  $N \times N$  patch in the HR Pan image. For all  $I \times J$  LR MS pixels, there are  $I \times J$  corresponding patches of size  $N \times N$ , and vice versa. Note that these  $I \times J$  patches do not overlap with each other and tile the whole Pan image. We learn a dictionary using those patches. To use that dictionary and form an HR MS image of the size  $NI \times NJ$  we then consider all of the overlapped patches of size  $N \times N$ , each mapping to an HR MS pixel. This mapping process allows us to use the learned dictionary on all the HR MS pixels.

The training data are built as follows. Let  $\mathbf{z}_{i,j}^{MS} = [z_{i,j}^{(1)} \ z_{i,j}^{(2)} \ \dots \ z_{i,j}^{(M)}]^T$  represent the  $M$  dimensional vector of color values of pixel  $(i, j)$  in the MS image with mean value  $\bar{z}_{i,j}^{MS}$ , and  $\mathbf{z}_{(Ni, Nj)}^{Pan} = [z_{(N(i-1)+1, N(j-1)+1)} \ \dots \ z_{(Ni, Nj)}]^T$  the  $N^2$  dimensional vector of gray values of the corresponding  $N \times N$  patch in the Pan image cornered at pixels  $(N(i-1)+1, N(j-1)+1)$  and  $(Ni, Nj)$ , with mean value  $\bar{z}_{i,j}^{Pan}$ . Dictionary learning algorithms typically subtract the mean value of each training data vector. Yet, our experience demonstrated that the mean values of the color spectrum and Pan patches are important for pan-sharpening. Thus, we keep the mean values as extra features in the training data vector. In other words, for each pixel in the MS image, we consider an  $M+1$  dimensional MS feature vector  $\mathbf{y}_{i,j}^{MS} = [(\mathbf{z}_{i,j}^{MS})^T - \bar{z}_{i,j}^{MS}, \bar{z}_{i,j}^{MS}]^T$ , with  $i = 1, 2, \dots, I$  and  $j = 1, 2, \dots, J$ , and for each corresponding Pan patch an  $N^2+1$  dimensional Pan feature vector  $\mathbf{y}_{i,j}^{Pan} = [(\mathbf{z}_{i,j}^{Pan})^T - \bar{z}_{i,j}^{Pan}, \bar{z}_{i,j}^{Pan}]^T$ . The training data incorporate spatial information by coupling the two

---

### Algorithm 1 Iterative algorithm to minimize (2)

---

- 1: **Input**  $\epsilon_0, K, \lambda$ , and  $T_0$
  - 2: **initialize**  $k = 0$ ,  $\text{err} = 1$ ,  $\mathbf{X}_0^{MS} = \bar{\mathbf{X}}^{MS}$ .
  - 3: **while**  $k < K$  and  $\text{err} > \epsilon_0$  **do**
  - 4:    $k = k + 1$
  - 5:    $\tilde{\alpha} = \operatorname{argmin}_{\alpha} \|\mathbf{X}_{k-1} - \tilde{\mathbf{D}}\alpha\|_F^2 \quad \text{s.t. } \forall l, \|\alpha_l\|_0 \leq T_0$
  - 6:    $\tilde{\mathbf{X}}_k = \tilde{\mathbf{D}}\tilde{\alpha}$
  - 7:    $\mathbf{X}_k^{MS} = \frac{\mathbf{E}^{MS}\tilde{\mathbf{X}}_k + \lambda\bar{\mathbf{X}}^{MS}}{1+\lambda}$
  - 8:    $\mathbf{X}_k = [(\mathbf{X}_k^{MS})^T (\mathbf{X}^{Pan})^T]^T$
  - 9:    $\text{err} = \frac{\|\mathbf{X}_k^{MS} - \mathbf{X}_{k-1}^{MS}\|_F}{\|\mathbf{X}_k^{MS}\|_F}$
  - 10: **end while**
  - 11: **Output**  $\hat{\mathbf{X}} = \mathbf{X}_k$ .
- 

feature vectors in  $\mathbf{y}_{i,j} = [(\mathbf{y}_{i,j}^{MS})^T (\mathbf{y}_{i,j}^{Pan})^T]^T$ . For all pixels in the MS image, we form an  $(M + N^2 + 2) \times (IJ)$  training data matrix  $\mathbf{Y} = \{\mathbf{y}_{i,j}\}$ . Note that the training data are built using only the Pan image and the MS image to be sharpened. We learn the dictionary using the K-SVD [7].

The testing set used to recover the missing data in the HR MS image are built in the similar way, except that the Pan patches are overlapped. Let  $\mathbf{x}_{p,q} = [(\mathbf{x}_{p,q}^{MS})^T (\mathbf{x}_{p,q}^{Pan})^T]^T$  represent the testing data vectors with  $p = 1, 2, \dots, NI$  and  $q = 1, 2, \dots, NJ$ , and  $\mathbf{X} = [(\mathbf{X}^{MS})^T (\mathbf{X}^{Pan})^T]^T$  the testing data matrix. Since we do not have HR MS image available, most of the MS related entries in the testing data are missing and are filled in using the learned dictionary.

## 2.3. Iterative updating algorithm

Given the learned dictionary, we need to predict the missing data in the testing set. One option could be to fill in the missing data using principles of DL-based image inpainting [5]. Unfortunately, this approach does not work well because the missing data are not randomly distributed in the testing data matrix but uniformly distributed in the MS image. Intuitively, since different colors may correspond to the same gray level, it is impossible to predict colors according to gray value of a pixel without knowing any other information.

To solve this problem, more constraints on the testing data are necessary. Considering the fact that the interpolated MS image is a good estimate of the HR MS image, we put a constraint on the distance between the interpolated MS image and the DL prediction. In other words, we treat the interpolated MS image as a noisy version of the HR MS image, with a constraint on the noise level. We formulate the objective function as follows.

$$\hat{\mathbf{X}} = \operatorname{argmin}_{\mathbf{X}, \alpha} \|\mathbf{X} - \tilde{\mathbf{D}}\alpha\|_F^2 + \lambda \|\mathbf{E}^{MS}\mathbf{X} - \bar{\mathbf{X}}^{MS}\|_F^2 \text{ s.t. } \forall l, \|\alpha_l\|_0 \leq T_0 \text{ and } \mathbf{E}^{Pan}\mathbf{X} = \mathbf{X}^{Pan}, \quad (2)$$

where  $\mathbf{E}^{MS}$  is the first  $(M+1)$  rows and  $\mathbf{E}^{Pan}$  is the last  $(N^2+1)$  rows of an  $(M+N^2+2) \times (M+N^2+2)$  identity matrix, and  $\bar{\mathbf{X}}^{MS}$  denotes the MS feature of the linear interpolated MS image. The weight  $\lambda$  controls the contribution of the interpolated MS image. When  $\lambda = \infty$ , the solution of (2) is the interpolated MS image; and when  $\lambda = 0$ , (2) is equivalent to the image inpainting model in [7]. Note that (2) is not a convex problem and a global optimal is not guaranteed. To solve (2), we develop the iterative algorithm in Algorithm 1.

Once we obtain  $\hat{\mathbf{X}}$ , we can map it onto the MS image  $\hat{\mathbf{Z}}$ . Since the training data do not include any HR MS image information, the MS image with dictionary learning is still a little bit blurry. Nevertheless, the DL based MS image resolution is much better than

the interpolated MS image, as observed in the experimental results shown in Section 3. Thus, color distortion is reduced when using one of the standard pan-sharpening methods in subsequent processing.

### 3. EXPERIMENTS

#### 3.1. Evaluation methods

To demonstrate the improvement of pan-sharpening with dictionary learning, we pan-sharpen the degraded images and compare the fused MS image with the source MS image before degradation (the ground truth). There are several metrics to quantitatively analyze the quality of pan-sharpened images; we consider the spectral angle mapper (SAM), the correlation coefficient (CC), and the signal-to-noise ratio (SNR).

The SAM is a global measure of spectral distortion. Let  $\mathbf{z}_{i,j} = [z_{i,j}^{(nir)} \ z_{i,j}^{(r)} \ z_{i,j}^{(g)} \ z_{i,j}^{(b)}]^T$  and  $\hat{\mathbf{z}}_{i,j} = [\hat{z}_{i,j}^{(nir)} \ \hat{z}_{i,j}^{(r)} \ \hat{z}_{i,j}^{(g)} \ \hat{z}_{i,j}^{(b)}]^T$  denote the spectral vectors of pixel  $(i, j)$  of the source image  $\mathbf{Z}$  and the fused image  $\hat{\mathbf{Z}}$ , respectively. The SAM is calculated by

$$\text{SAM} = \sum_{i=1}^I \sum_{j=1}^J \arccos \left( \frac{\langle \mathbf{z}_{i,j}, \hat{\mathbf{z}}_{i,j} \rangle}{\|\mathbf{z}_{i,j}\|_2 \cdot \|\hat{\mathbf{z}}_{i,j}\|_2} \right). \quad (3)$$

If the fused image is identical to the source image, the SAM value goes to zero.

The correlation coefficient indicates the degree of correlation between  $\mathbf{Z}$  and  $\hat{\mathbf{Z}}$ . The correlation coefficient (CC) of band  $b$  is calculated by

$$\text{CC}^{(b)} = \frac{\sum_i \sum_j [z_{i,j}^{(b)} - \bar{z}^{(b)}][\hat{z}_{i,j}^{(b)} - \bar{\hat{z}}^{(b)}]}{\sqrt{\sum_i \sum_j [z_{i,j}^{(b)} - \bar{z}^{(b)}]^2 \sum_i \sum_j [\hat{z}_{i,j}^{(b)} - \bar{\hat{z}}^{(b)}]^2}}, \quad (4)$$

where  $z_{i,j}^{(b)}$  and  $\hat{z}_{i,j}^{(b)}$  denote the pixel values of  $(i, j)$  in band  $b$  of the source MS image and the corresponding fused image respectively, and  $\bar{z}^{(b)}$  and  $\bar{\hat{z}}^{(b)}$  denote the overall average values of band  $b$  respectively. When  $\hat{\mathbf{Z}} = \mathbf{Z}$ , the correlation coefficient equals one.

The signal-to-noise ratio (SNR) gives the standard measure of noise level. It is defined in dB as

$$\text{SNR}^{(b)} = 10 \log_{10} \frac{\sum_{i=1}^I \sum_{j=1}^J (z_{i,j}^{(b)})^2}{\sum_{i=1}^I \sum_{j=1}^J (z_{i,j}^{(b)} - \hat{z}_{i,j}^{(b)})^2}. \quad (5)$$

The higher the SNR, the better.

#### 3.2. QuickBird Data

The QuickBird satellite system provides Pan images at 0.7m resolution and MS images at 2.8m resolution with four bands of near infra-red (NIR), red, green, and blue [9]. In order to evaluate our pan-sharpening results we manually degrade the Pan and MS images with low-pass filter (Gaussian kernel with unit variance) and downsample by four to yield one 2.8m Pan image and one 11.2m resolution MS image. Pan-sharpening is then performed on the degraded images and the pan-sharpened results are compared with the source 2.8m MS images. Fig. 1 shows an example of the degraded QuickBird MS image and Pan image of Sundarbans. For the MS image we only plot the RGB components, although we include the NIR component in our evaluation. The degraded Pan image is composed of  $320 \times 320$  pixels and the degraded MS image  $80 \times 80$  pixels, both covering an identical area of  $896\text{m} \times 896\text{m}$ . Given the degraded Pan and MS images, we learn an over-complete dictionary of 300



(a) Low resolution MS image

(b) Pan image

Fig. 1. Downgraded QuickBird MS and Pan image



(a) DL based MS image

(b) Interpolated MS image

Fig. 2. MS image using (a) DL, and (b) Bicubic interpolation.

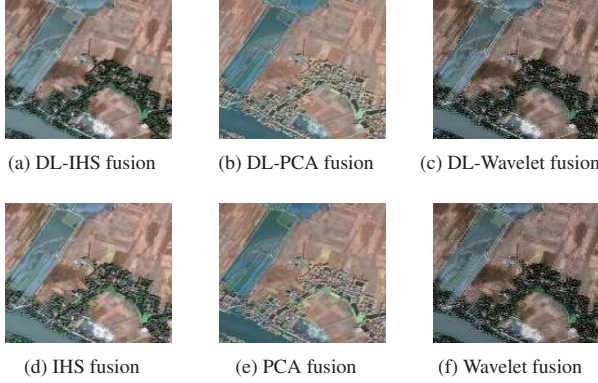
atoms. With the learned dictionary, we update the MS image using the proposed iterative algorithm with parameter settings  $\epsilon_0 = 10^{-3}$ ,  $K = 20$ ,  $\lambda = 1$ . The results are shown in Fig. 2(a). For comparison, we plot the classical bicubic interpolated MS image in Fig. 2(b). We observe that the MS image resolution is significantly improved with dictionary learning, especially in the texture edges. For further improvement, we consider IHS, PCA, and wavelet based pan-sharpening, with the corresponding results shown in Fig. 3. The pan-sharpening results using DL are noticeably better than those without DL. To quantify the improvement, the SAM, CC, SNR values are computed as listed in Table 1. All values in the table indicate that the pan-sharpening performance is improved using the DL process.

#### 3.3. IKONOS Data

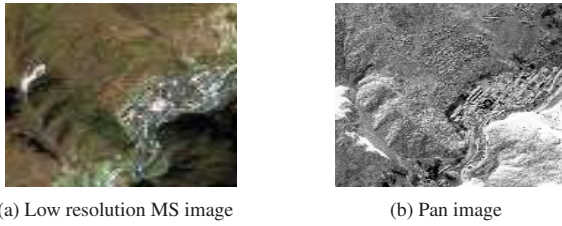
The IKONOS satellite captures 1m-resolution Pan images and 4m-resolution MS image covering NIR, red, green, and blue bands as well [10]. Similarly, we perform pan-sharpening on degraded images and compare the results with the original MS images. Fig. 4 shows an example of the degraded IKONOS images (RGB and Pan) of Sichuan, China. The degraded Pan image is  $320 \times 320$  pixels and the degraded MS image is  $80 \times 80$  pixels, both covering an area of  $320\text{m} \times 320\text{m}$ . Again, we learn a dictionary of 300 atoms, and test on the testing data. The MS image result after DL is shown in Fig. 5(a) with improved resolution, compared to the interpolated MS image is

Table 1. Comparison of fusion results with degraded QuickBird data

	DL-IHS	IHS	DL-PCA	PCA	DL-wavelet	Wavelet	
SAM	<b>.1066</b>	.1119	<b>.1719</b>	.1994	<b>.1412</b>	.1445	
CC	nir	<b>.9625</b>	.9301	<b>.9445</b>	.7759	<b>.9556</b>	.9156
	r	<b>.9338</b>	.9112	<b>.5237</b>	.5191	<b>.8995</b>	.8857
	g	<b>.9190</b>	.8871	<b>.4969</b>	.4846	<b>.8706</b>	.8559
	b	<b>.9097</b>	.8677	<b>.5474</b>	.4738	<b>.8515</b>	.8368
SNR	nir	<b>17.97</b>	14.90	<b>14.93</b>	10.21	<b>17.53</b>	14.79
	r	<b>15.98</b>	14.40	<b>8.24</b>	8.07	<b>14.22</b>	13.68
	g	<b>15.84</b>	14.15	<b>9.15</b>	9.02	<b>14.09</b>	13.60
	b	<b>14.96</b>	13.15	<b>8.83</b>	8.41	<b>12.81</b>	12.41
	all	<b>16.00</b>	14.10	<b>9.44</b>	8.77	<b>14.27</b>	13.51



**Fig. 3.** Pan-sharpened QuickBird image using (a) DL-IHS, (b) DL-PCA, (c) DL-Wavelet, (d) IHS, (e) PCA, and (f) Wavelet.



**Fig. 4.** Downgraded IKONOS Sichuan MS and Pan image

shown in Fig. 5(b). Using these MS images, we perform IHS, PCA, and wavelet based pan-sharpening with the pan-sharpened results shown in Fig. 6 and the SAM, CC, SNR values listed in Table 2. We observed consistent improvement with color distortion, as indicated by the SAM, while for the CC and SNR, most values are improved, with few exceptions.

#### 4. CONCLUSION

We propose a dictionary learning-based pan-sharpening process. Our experimental results show the improvement using dictionary learning in a number of figures of merit, including the spectral angle mapper (SAM), the correlation coefficient(CC) and the signal-to-noise ratio (SNR), as well as in subjective evaluation. Compared with other existing pan-sharpening algorithms—including dictionary-learning based ones—our dictionary learning process does not require a large amount of high resolution MS images for training, and can be incorporated in other pan-sharpening algorithms

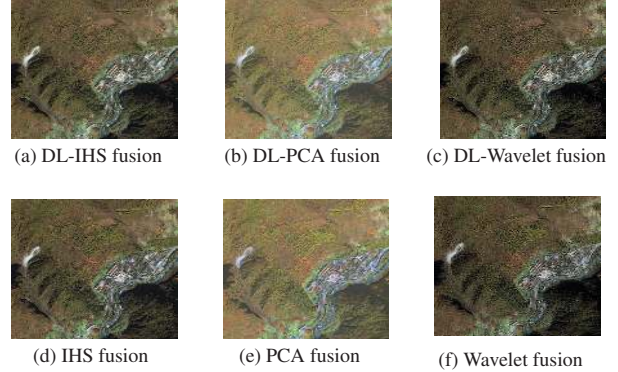
**Table 2.** Comparison of fusion results with degraded IKONOS data

	DL-IHS	IHS	DL-PCA	PCA	DL-wavelet	Wavelet
SAM	<b>.1203</b>	.1293	<b>.1176</b>	.1369	<b>.1581</b>	.1617
nir	<b>.9543</b>	.9496	<b>.8632</b>	.8182	<b>.9475</b>	.9195
r	.9134	<b>.9149</b>	<b>.8759</b>	.8496	<b>.8981</b>	.8756
g	.9215	<b>.9245</b>	<b>.9065</b>	.8828	<b>.9052</b>	.8838
b	<b>.8761</b>	.8683	<b>.8835</b>	.8286	<b>.8424</b>	.8260
nir	<b>17.47</b>	16.75	<b>12.55</b>	11.58	<b>17.41</b>	15.59
r	<b>14.54</b>	14.52	<b>10.76</b>	10.56	<b>13.92</b>	13.05
g	14.65	<b>14.68</b>	<b>10.71</b>	10.45	<b>13.85</b>	13.00
b	<b>11.76</b>	11.53	<b>9.70</b>	9.21	<b>10.36</b>	9.87
all	<b>14.88</b>	14.66	<b>11.18</b>	10.68	<b>14.10</b>	13.16



(a) DL based MS image (b) Interpolated MS image

**Fig. 5.** MS image using (a) DL, and (b) Bicubic interpolation.



**Fig. 6.** Pan-sharpened IKONOS image using algorithms (a) DL-IHS, (b) DL-PCA, (c) DL-Wavelet, (d) IHS, (e) PCA, and (f) Wavelet.

that use MS interpolation. Our dictionary learning process can be easily extended to fuse hyper-spectral data with small modifications.

#### 5. REFERENCES

- [1] Yun Zhang, "Understanding image fusion," *Photogrammetric Engineering and Remote Sensing*, pp. 657–661, June 2004.
- [2] Luciano Alparone, Lucien Wald, Jocelyn Chanussot, Claire Thomas, Paolo Gamba, and Lori Mann Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. remote sens.*, vol. 45(10), pp. 3012–3021, October 2007.
- [3] Qian Du, Oguz Gungor, and Jie Shan, "Performance evaluation for pan-sharpening techniques," in *IEEE Geosci. Remote Sens. Symposium*, 2005, pp. 4264–4266.
- [4] Claire Thomas, Thierry Ranchin, Lucien Wald, and Jocelyn Chanussot, "Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46(5), pp. 1301–1312, May 2008.
- [5] Atul Divekar and Okan Ersoy, "Image fusion by compressive sensing," in *Geoinformatics, 2009 17th International Conference on*, 2009.
- [6] Shutao Li and Bin Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. remote sens.*, vol. 49(2), pp. 738–746, February 2011.
- [7] Michal Aharon, Michael Elad, and Alfred Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Sig. Proc.*, vol. 54(11), pp. 4311–4322, 2006.
- [8] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma, "Image super-resolution via sparse representation," *IEEE Trans. Im. Proc.*, vol. 19(11), 2010.
- [9] DigitalGlobe, "QuickBird data repository," available at <http://www.glcf.umiacs.umd.edu/data/quickbird/>.
- [10] DigitalGlobe, "IKONOS data repository," available at <http://www.glcf.umiacs.umd.edu/data/ikonos/>.