21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Denoising Hyperspectral Images Using Spectral Domain Statistics

Antony Lam and Imari Sato
Digital Content and Media Sciences Research Division
National Institute of Informatics
antony@nii.ac.jp; imarik@nii.ac.jp

Yoichi Sato
Institute of Industrial Science
The University of Tokyo
ysato@iis.u-tokyo.ac.jp

Abstract

Hyperspectral imaging has proven useful in a diverse range of applications in agriculture, diagnostic medicine, and surveillance to name a few. However, conventional hyperspectral images (HSIs) tend to be noisy due to limited light in individual bands; thus making denoising necessary. Previous methods for HSI denoising have viewed the entire HSI as a general 3D volume or focused on processing the spatial domain. However, past findings suggest that spectral distributions exhibit less variation than spatial patterns. Hence it would be fruitful to take specific advantage of the more predictable behavior of spectral domain data for denoising. In this paper, we present a two-stage denoising framework that first emphasizes denoising in the spectral domain and then uses spatial information to further improve spectral domain denoising. Our results indicate that specifically leveraging the spectral domain for denoising can provide state-of-the-art performance even from a relatively simple approach.

1 Introduction

Hyperspectral imaging is the process of capturing images of a scene over multiple bands of the electromagnetic spectrum. When captured, a hyperspectral image (HSI) can be thought of as a set of 2D spatially organized "pixels" where each pixel contains an entire spectral distribution over wavelengths. This allows us to see the spectral distribution of any given surface point in a scene and has lead to numerous applications in agriculture, diagnostic medicine, surveillance, and more. However, conventional hyperspectral imaging suffers from limited light in individual bands which introduces noise into the imaging process.

In this paper, we present a simple but effective denoising method that exploits the spectral domain statistics present in all HSIs. This is in contrast to past approaches that treat HSIs as generic 3D volumes or focus primarily on the spatial domain. For example, in [12], the HSI was treated as a generic 3D volume. The algorithm is general but performance is limited since specific HSI characteristics are not taken advantage of. Later approaches [2, 4, 7] independently process homogeneous spatial subregions. However, this approach is computationally expensive and does not consider global scene statistics. Also, the typical use of rectangular spatial subregions may not be ideally suited to all scenes. We note that methods such as [13] produce arbitrarily shaped subregions. However, specific HSI characteristics are also not leveraged.

(a) Noisy (SNR 1.38) (b) Denoised (SNR 20.23)

Figure 1: Sample denoising result at 500 nm.

Other methods [1, 3, 10] process HSIs globally while treating the spatial and spectral domain as separable. Specifically, [1] decorrelates the spectral and spatial domains and then denoises via wavelet thresholding. A drawback is that the noise variance for each decorrelated spectral band needs to be estimated. In both [3] and [10], the spatial and spectral domains are denoised separately. Their common drawback is that each of the spectra are denoised independently so spectral domain statistics are not leveraged. Indeed, the past and recent methods discussed do not take full advantage of the wealth of information in the spectral domain.

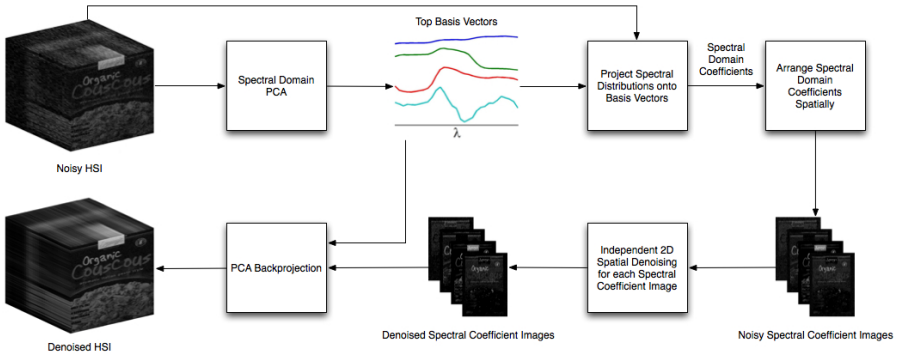

Figure 2: Overview of our HSI Denoising Framework

2 Denoising with the Spectral Domain

Our emphasis on the spectral domain is motivated by the fact that the spectral domain exhibits much less variation in terms of patterns than the spatial domain. To see why, we first note it is well known that large sets of spectra consisting of natural scenes or extensive color samples can be represented compactly in just 6-8 principal components [6, 8, 9, 11]. On the other hand, a simple test on the UEA Multispectral Dataset [5] reveals that 1,260 random 10x10 spatial patches require 29 principal components to capture 99% of the variance. The requirement of a much higher number of principal components for spatial patches indicates the more complex nature of the spatial domain. Thus our intuition is that noise is more easily separated from the less varied and better understood patterns in the spectral domain.

Having noted the complex nature of spatial patches, we point out that the spatial domain of a *single HSI* can be more compactly represented if entire band images are treated as individual patches.[1] In that case, we have found that only six principal components are typically required. However, we will show that even this approach (spatial PCA) to analyzing HSIs still underperforms in comparison to our spectral domain approach.

Another advantage of the spectral domain is that typical HSIs contain an overwhelming number of samples in the spectral domain. For example, a 256x256 scene would contain 65,536 spectra. We find this makes it possible to robustly separate true signals from noise by only analyzing the noisy image in question.

[1] We found that 100 random 10x10 patches of a single HSI, would still require a high number of principal components.

2.1 Overview

We now present an overview of our method. We aim to fully exploit the favorable conditions of the spectral domain while incorporating spatial information. This is accomplished by first building a low-noise spectral domain basis from the noisy HSI via PCA. The noisy HSI's spectra are then projected onto the basis to obtain coefficients. We then additionally denoise the coefficients using spatial information. The HSI can then be recovered via backprojection. See Fig. 2 for an visualization of our framework.

2.2 Low-Noise Basis from High-Noise Images

One of the main advantages of the spectral domain over the spatial domain is the computation of low-noise basis vectors even from very noisy HSIs. This is because of the two main characteristics of spectral domain data mentioned earlier. Spectral distributions can be very compactly represented and there is an overwhelmingly large number of spectra (e.g. over 65,000) in a typical HSI. As we also expect noise to occur randomly, the impact of noise on the greatest directions of variance in the large set of spectra in an HSI should be minimal. Thus PCA can be used to obtain low-noise top basis vectors even when noise is high.[2] Using the low-noise basis, one could arrange all the noisy HSI's spectra as a column vector $\mathbf{R}$ and denoise as follows:

$$\mathbf{R_D} = \mathbf{B}\mathbf{B}^{\mathbf{T}}(\mathbf{R} - \mathbf{M}) + \mathbf{M} = \mathbf{B}\mathbf{C} + \mathbf{M} \quad (1)$$

where $\mathbf{B}$ is the column matrix of basis vectors, $\mathbf{M}$ is a matrix with each column consisting of the HSI's empirical mean spectral distribution, and $\mathbf{C}$ are the basis coefficients. $\mathbf{R_D}$ is the matrix of denoised spectra which can be rearranged to recover the HSI band images.

[2] Validation of our claim is shown in the experiments.

2.3 Incorporating Spatial Information and Recovering the Denoised Image

Projection onto the low-noise basis removes significant noise. However, noise would still be present in coefficients $\mathbf{C}$ as some noisy spectra can exist in the space spanned by the basis. Fortunately, spatial information can be used to denoise coefficients. This is because the spatial relationship between scene points carry over to the spatial relationships between the points' spectral domain coefficients. Thus spatially arranging coefficients into a set of 2D spectral coefficient images (SCIs)[3] exhibit spatial components of the scene. This suggests that each SCI could be independently denoised using off-the-shelf 2D spatial denoisers.

However, is it reasonable to independently process each SCI? The nature of PCA indicates the answer is yes. PCA produces basis vectors that are orthogonal to each other. So changes in one SCI do not affect the other components of the spectra. Thus each SCI can be processed independently while still enforcing constraints over all wavelengths.

In our paper, we chose to use the bilateral filter [14] to denoise SCIs. Briefly, the bilateral filter is applied to each coefficient at location (x,y) in SCI S as

$$S_D(x,y) = k_{x,y}^{-1} \sum_{(i,j)\in N} S(i,j)\Big(G_s\big(i-x, j-y\big) G_r\big(S(i,j) - S(x,y)\big)\Big) \quad (2)$$

where $k_{x,y}$ is a normalization factor, N is the neighborhood of pixels around and including (x,y), G_s is a 2D Gaussian function centered at (0,0) with standard deviation σ_s for both dimensions, and G_r is a 1D Gaussian function centered at 0 with standard deviation σ_r. Due to the spatial structure in SCIs, we have found the edge preserving capabilities of the bilateral filter provides excellent denoising of coefficients. Once denoised, the coefficients can be used in Eq. 1 to fully denoise the HSI.

3 Experiments

3.1 Experimental Setup

We use the UEA Multispectral Image Database [5] for our tests. This database contains images of everyday objects captured from 400 - 700 nm in increments of 10 nm under CIE illuminant D75. In our experiments, we use 20 objects from the database.

[3]One SCI for each basis vector.

For our first set of experiments on comparing the computation of basis vectors in the spectral versus the spatial domain, we use the percent error between two vectors $\vec{a}$ and $\vec{b}$ defined as

$$P_{error} = \left(\frac{\|\vec{a} - \vec{b}\|}{\|\vec{a}\|}\right) \times 100\% \quad (3)$$

where $\|\cdot\|$ denotes the norm of the vector.

As in past HSI denoising work [3], we compute SNR as

$$SNR = \frac{\sum_{i,j,k} A(i,j,k)^2}{\sum_{i,j,k}[A(i,j,k) - B(i,j,k)]^2} \quad (4)$$

where A is the clean HSI and B is the noisy HSI.

As is commonly done in the literature, we added white Gaussian noise to our HSIs. For noise variance, we set σ values to 3%, 5%, 10%, and 20% of the dynamic range of each HSI being noised.

3.2 Methods Compared Against

Our method is compared against denoising using spatial PCA and spectral PCA. We also compared against Chen and Qian (CQ) [3] and Atkinson et al. (AKJ) [1]. Both CQ and AKJ work with the spatial and spectral domains but do not take full advantage of spectral domain statistics. We also tried spectral PCA followed by denoising each band image independently. [4] We call the last method Spectral PCA + Independent Band Denoising (IBD).

3.3 Experimental Results

Spatial Domain Eigenvector #						
Ave SNR	1	2	3	4	5	6
47.2	3%	8%	18%	26%	29%	44%
17.0	15%	70%	126%	121%	133%	127%
4.3	29%	122%	140%	135%	133%	129%
1.0	59%	147%	135%	137%	140%	138%

Spectral Domain Eigenvector #				
Ave SNR	1	2	3	4
47.2	0.03%	0.07%	0.17%	11%
17.0	0.1%	0.4%	1%	14%
4.3	0.3%	0.8%	2.5%	44%
1.0	0.6%	22%	27%	94%

Table 1: Mean Percent Errors Between Eigenvectors from Noisy and Clean HSIs

[4]We note that the bilateral filter is used in both independent band image denoising and SCI denoising where the window size is set to be 3x3, the spatial sigma is 1x1, and the range sigma is 5% of the dynamic range of all the SCIs.

Noisy	Ours (4)	Spatial PCA (3)	Spatial PCA (6)	Spectral PCA (4)	CQ (3)	AKJ	Spectral PCA (4) + IBD
47.2	203.5	146.1	166.7	159.6	145.4	46.8	**211.8**
17.0	**165.6**	84.0	68.8	88.6	83.9	17.0	149.3
4.3	**69.5**	28.7	18.3	29.2	28.7	4.3	50.8
1.0	**13.2**	7.9	4.6	7.9	7.9	1.2	11.3

Table 2: Mean SNR results. Parenthesis show number of top eigenvectors. IBD means Independent Band Denoising.

Figure 3: Sample denoising result using our method. The 400 nm band is shown.

We first present results to validate our claim in Sec. 2.2. Table 1 compares the similarity between eigenvectors computed from noisy HSIs and counterparts from the noise-free case. The number of eigenvectors were chosen to capture 99% of the variance in the noise-free case. The low errors for the spectral domain eigenvectors indicates the low amount of noise in them. Also note that error rates only increase for lower ranked eigenvectors. This is fortunate as lower ranked eigenvectors are less important in data reconstruction.

Table 2 shows a comparison of our method against others. For CQ, the best number of eigenvectors was determined manually. Our use of the spectral domain resulted in the best performance. Interestingly, Spectral PCA + IBD performed slightly better than our method in the first row. This is likely because the benefits of our enforcing joint constraints over wavelengths is less important at low noise. However, Spectral PCA + IBD is slow since all 31 band images are processed rather than four SCIs in our method. A visual result of our method from Table 2's second row is in Fig 3.

4 Conclusion

By taking advantage of the well-behaved nature of spectra and the overwhelmingly large number of spectra found in HSIs, we were able to show robust performance over a wide range of noise levels. In the future, we hope to explore other decompositions besides PCA and also non-local approaches to SCI denoising.

References

[1] I. Atkinson, F. Kamalabadi, S. Mohan, and D. Jones. Wavelet-based 2-d multichannel signal estimation. In *ICIP*, volume 2, pages II – 141–4 vol.3, sept. 2003. 1, 3

[2] S. Bourennane, C. Fossati, and A. Cailly. Improvement of target-detection algorithms based on adaptive three-dimensional filtering. *IEEE Trans. Geosci. Remote Sens.*, 49(4):1383–1395, 2011. 1

[3] G. Chen and S.-E. Qian. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Trans. Geosci. Remote Sens.*, 49(3):973–980, 2011. 1, 3

[4] J.-M. Gaucel, G. Mireille, and S. Bourennane. Adaptive linear minimum mean square error restoration: influence on hyperspectral detection strategy. *Int. J. Remote Sens.*, 29:2943–2961, January 2008. 1

[5] S. Hordley, G. Finalyson, and P. Morovic. A multispectral image database and its application to image rendering across illumination. In *ICIG*, pages 394–397, Washington, DC, USA, 2004. IEEE Computer Society. 2, 3

[6] T.-W. Lee, T. Wachtler, and T. J. Sejnowski. The spectral independent components of natural scenes. In *Workshop on Biologically Motivated Computer Vision*, BMVC '00, pages 527–534, London, UK, 2000. Springer-Verlag. 2

[7] D. Letexier and S. Bourennane. Noise removal from hyperspectral images by multidimensional filtering. *IEEE Trans. Geosci. Remote Sens.*, 46(7):2061–2069, 2008. 1

[8] L. T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA A*, 3(10):1673–1683, Oct 1986. 2

[9] D. H. Marimont and B. A. Wandell. Linear models of surface and illuminant spectra. *JOSA A*, 9(11):1905–1913, Nov 1992. 2

[10] H. Othman and S.-E. Qian. Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage. *IEEE Trans. Geosci. Remote Sens.*, 44(2):397–408, 2006. 1

[11] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of munsell colors. *JOSA A*, 6:318–322, 1989. 2

[12] N. Renard, S. Bourennane, and J. Blanc-Talon. Multiway filtering applied on hyperspectral images. In *ACIVS*, pages 127–137, Berlin, Heidelberg, 2006. Springer-Verlag. 1

[13] P. Soille. Constrained connectivity and connected filters. *IEEE Trans. PAMI*, 30(7):1132–1145, July 2008. 1

[14] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, pages 839–846, Washington, DC, USA, 1998. IEEE Computer Society. 3