

# Statistics of Real-World Hyperspectral Images

Ayan Chakrabarti and Todd Zickler

Harvard School of Engineering and Applied Sciences

33 Oxford St, Cambridge, MA 02138.

{ayan, zickler}@eecs.harvard.edu

## Abstract

*Hyperspectral images provide higher spectral resolution than typical RGB images by including per-pixel irradiance measurements in a number of narrow bands of wavelength in the visible spectrum. The additional spectral resolution may be useful for many visual tasks, including segmentation, recognition, and relighting. Vision systems that seek to capture and exploit hyperspectral data should benefit from statistical models of natural hyperspectral images, but at present, relatively little is known about their structure. Using a new collection of fifty hyperspectral images of indoor and outdoor scenes, we derive an optimized “spatio-spectral basis” for representing hyperspectral image patches, and explore statistical models for the coefficients in this basis.*

## 1. Introduction

Most cameras capture three spectral measurements (red, green, blue) to match human trichromacy, but there is additional information in the visible spectrum that can be exploited by vision systems. Hyperspectral images, meaning those that provide a dense spectral sampling at each pixel, have proven useful in many domains, including remote sensing [2, 3, 5, 22, 35], medical diagnosis [10, 29, 33], and biometrics [31], and it seems likely that they can simplify the analysis of everyday scenes as well.

When developing vision systems that acquire and exploit hyperspectral imagery, we can benefit from knowledge of the underlying statistical structure. By modeling the interdependencies that exist in the joint spatio-spectral domain, we should be able to build, for example, more efficient systems for capturing hyperspectral images and videos, and perhaps better tools for visual tasks such as segmentation and recognition.

This paper seeks to establish the basic statistical structure of hyperspectral images of “real-world” scenes, such as offices, streetscapes, and parks, that we encounter in everyday life. Unlike previous analyses, which have sep-

arately considered the spectral statistics of point samples [18, 21, 25], we consider the spatial and hyperspectral dimensions jointly to uncover additional structure. Using a new collection of fifty hyperspectral images captured with a time-multiplexed 31-channel camera, we evaluate difference choices of spatio-spectral bases for representing hyperspectral image patches and find that a separable basis is appropriate. Then, we characterize the statistical properties of the coefficients in this basis and describe models that capture these properties effectively.

## 2. Related Work

Our work is motivated by successes in analyzing and modeling the statistical properties of grayscale images [1, 30, 39]. These models have proved valuable for inferring accurate images from noisy and incomplete measurements, with applications in denoising [12, 28] and restoration [4, 20]. These low-level statistics have also found use as building blocks for higher-level visual tasks such as segmentation and object detection [7, 23, 36]. Our work is also motivated by studies of the joint spatial-color structure of trichromatic images (corresponding to human cone responses, or the standard RGB color space) [15, 19, 27, 32, 38], which may have implications for tasks such as demosaicking for efficient RGB image capture [14, 16] and computational color constancy [6, 37]. Our goal in this paper is to develop models that are even more powerful by considering hyperspectral data and by considering the joint statistics of variations with respect to space and wavelength.

The present study is enabled by recent advances in hyperspectral capture systems, which include those based on spatial-multiplexing with generalized color filter arrays [41], spatial-multiplexing with a prism [11], time-multiplexing with liquid crystal tunable filters [13, 17], and time-multiplexing with varying illumination [24, 26]. Prior to these advances, studies of real-world spectra have been limited to collections of point samples, such as those collected by a spectrometer or spectroradiometer. These studies have suggested, for example, that the spectral reflectances of “real-world” materials are smooth functions

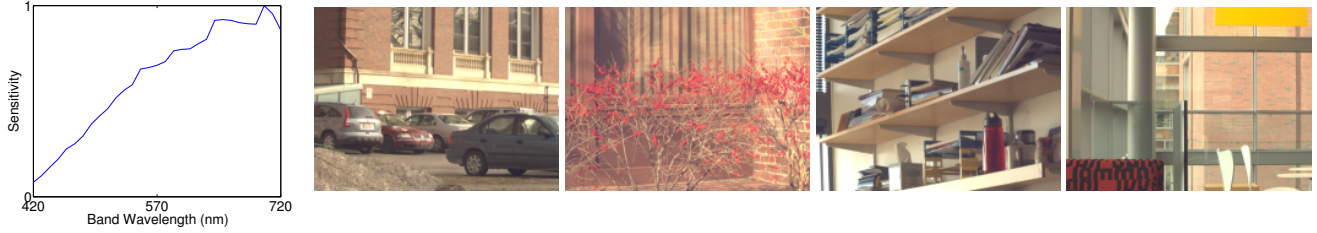


Figure 1. Hyperspectral Database of “real-world” images. Each image has a spatial resolution of  $1392 \times 1040$  with thirty-one spectral measurements at each pixel. Left: Camera sensitivity for each wavelength band. Right: Typical images of indoor and outdoor scenes from the database, rendered in sRGB.

that can be represented with 6-8 principal components [19, 21, 25] (or a suitable sparse code [18]), and that the spectra of daylight and other natural illuminants can be represented with even fewer principal components [40]. Our goal in the present study is to move beyond point samples, and to investigate variations in spectral distributions within spatial neighborhoods.

We expect that accurate statistical models will aid in the design of efficient hyperspectral acquisition systems. Many proposed acquisition methods seek to reconstruct full spectral images from a reduced set of measurements based on assumptions about the underlying statistics [26, 41]. Such methods are likely to benefit from accurate statistical models that are learned from real-world hyperspectral data. These models may also prove useful for other applications, such as relighting, segmentation, and recognition.

Other hyperspectral datasets that are related to that introduced here include those of Hordley *et al.* [17] and Yasuma *et al.* [41]. These datasets include 22 and 32 hyperspectral images, respectively, and they are focused on objects captured with controlled illuminants in laboratory environments. More related is the database of 25 hyperspectral images of outdoor urban and rural scenes captured by Foster *et al.* [13]. A primary aim of our work has been to capture and analyze a larger database that includes both indoor and outdoor scenes.

### 3. Hyperspectral Image Database

To enable an empirical analysis of the joint spatio-spectral statistics of real-world hyperspectral scenes, we collected a database of fifty images under daylight illumination, both outdoors and indoors, using a commercial hyperspectral camera (Nuance FX, CRI Inc.) The camera uses an integrated liquid crystal tunable filter and is capable of acquiring a hyperspectral image by sequentially tuning the filter through a series of thirty-one narrow wavelength bands, each with approximately 10nm bandwidth and centered at steps of 10nm from 420nm to 720nm. Figure 1 (left) shows the relative sensitivity of the camera for each wavelength band, accounting for both the quantum-efficiency of the 12-bit grayscale sensor and the per-band transmittance

of the effective filters. The camera is equipped with an apochromatic lens (CoastalOpt UV-VIS-IR 60mm Apo Macro, Jenoptik Optical Systems, Inc.) and in all cases we used the smallest viable aperture setting. The combination of the apo-chromatic lens and the avoidance of a mechanical filter wheel allows us to acquire images that are largely void of chromatic aberration and mis-alignment. To avoid contaminating the statistics by having different per-band noise levels, we did not vary the exposure time across bands or normalize the captured bands with respect to sensitivity. All results in the following sections must therefore be interpreted relative to the camera sensitivity function. However, the appendix includes a discussion on statistics computed after normalizing for the sensitivity.

Due to the use of small apertures and the low transmittance of individual bands, the total acquisition times for an entire image (*i.e.*, all wavelength bands) are high and vary from fifteen seconds to over a minute. Accordingly, all images were captured using a tripod and by ensuring minimal movement in the scene. In the interest of having a diverse dataset, we have captured images with movement in some regions— but these regions (and other areas affected by dust, *etc.*) are masked out manually before analysis. We note that as a result, any regions with people in the captured scenes are masked out, and our analysis does not include samples of human skin tones.

The captured dataset includes images of both indoor and outdoor scenes featuring a diversity of objects, materials and scale (see Fig. 1 for a few example images rendered in sRGB). We believe the database to be a representative sample of real-world images, capturing both pixel-level material statistics and spatial interactions induced by texture and shading effects. In addition to the analysis here, these images may be useful “ground truth” to design and evaluate methods for acquisition and vision tasks. We have also captured twenty-five additional images taken under artificial and mixed illumination, and while these are not used for the analysis presented in this work, they are being made available to the community along with the fifty natural illumination images. The entire database is available for download at <http://vision.seas.harvard.edu/hyperspec/>.

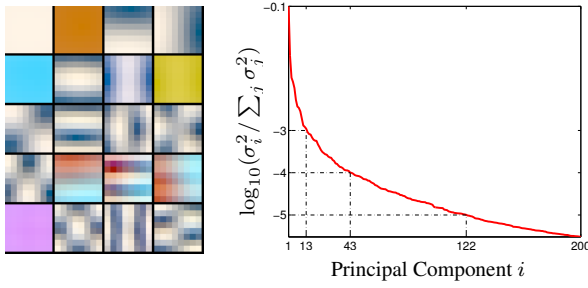


Figure 2. General basis for  $8 \times 8$  hyper-spectral patches learned across the database. Left: most significant basis vectors (in reading order) rendered in RGB. Right: variance of the coefficients for the first 200 basis vectors. The variance decays rapidly indicating that a small proportion of components are sufficient for accurate reconstruction.

## 4. Spatio-Spectral Representation

In this section, we explore efficient representations for hyperspectral images. As is common practice with grayscale and RGB images, we first divide the entire image into patches and consider the properties of each patch independently. Let  $X[n, l]$  be a random  $P \times P$  hyperspectral image patch, where  $n \in \{1, \dots, P\}^2$  and  $l \in \{1, \dots, 31\}$  index pixel location and spectral band respectively. For the rest of the paper, we choose the patch size  $P = 8$ , but the results and conclusions from different choices of  $P$  are qualitatively the same.

Since  $X$  is high-dimensional, we seek a representation that allows analysis in terms of a smaller number of components. Formally, we wish to find an optimal orthonormal basis set  $\{V_i\}$  and express  $X$  in terms of scalar coefficients  $x_i$  as

$$X[n, l] = \mu[n, l] + \sum_i x_i V_i[n, l], \quad (1)$$

where  $x_i = \langle X - \mu, V_i \rangle$ , and  $\mu$  is the “mean patch”.

We begin by learning a set of general basis vectors using principal component analysis (PCA) on patches cropped from images in the database. Figure 2 shows the top twenty components rendered in RGB, as well as the variance for the top 200 components. We see that the first two  $V_i$  essentially correspond to spatially-constant “DC” components with distinct spectral variation, followed by vertical and horizontal derivative components. We also find that there is a sharp fall off in variance indicating that  $X$  can be described accurately by a relatively small number of coefficients. Indeed, the first 20 basis vectors (out of a total of nearly 2000) account for 99% of the total variance.

### 4.1. Separable Basis Components

We observe that the basis set in Fig. 2 has sets of vectors with similar spatial patterns but different spectra or “colors”. Therefore, we explore the utility of a “separable” basis

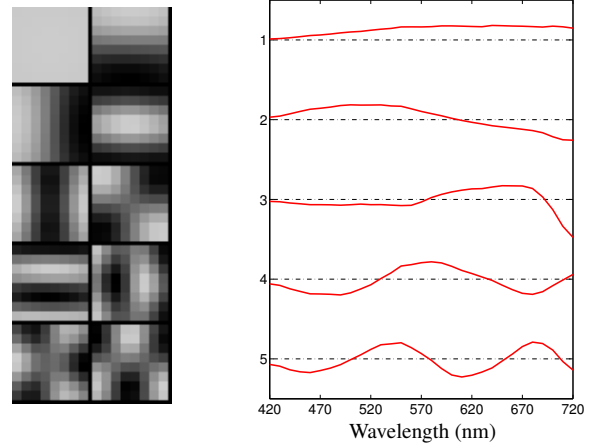


Figure 3. Learned separable basis, with most significant spatial components  $\{S_j\}$  (left) and spectral components  $\{C_k\}$  (right). The overall basis vectors for a patch are then given by  $S_j[n]C_k[l]$ , for every pair  $(j, k)$ .

set, meaning one in which every  $V_i$  can be decomposed into a Cartesian product of separate spatial and spectral components. Notationally, we write  $V_i[n, l] = S_j[n]C_k[l]$ , where  $\{S_j\}_{j=1}^{P^2}$  and  $\{C_k\}_{k=1}^{31}$  are orthonormal bases spanning the space of monochrome  $P \times P$  spatial patches and the space of 31-channel spectral distributions, respectively. Note that by construction the components  $\{V_i\}$  formed by different combinations of  $S_j$  and  $C_k$  also form an orthonormal basis. Again, we use PCA to learn  $\{S_j\}_j$  from monochrome patches pooled across all bands and  $\{C_k\}_k$  from the spectral distributions at individual pixels. The combined spatio-spectral basis vectors are then formed as  $S_j[n]C_k[l]$ , for all pairs of  $j$  and  $k$ .

Figure 3 shows the first few spatial and spectral components,  $\{S_j\}$  and  $\{C_k\}$ , learned in this manner. The spatial components correspond to a DCT-like basis used commonly for modeling grayscale images, with  $S_1$  corresponding to the “DC” component. The spectral components in turn resemble a Fourier basis scaled by the camera’s sensitivity function (see Fig. 1).

In Fig. 4, we compare this learned separable basis to the one derived with general PCA in terms of the relative reconstruction error

$$\epsilon(i) = \log_{10}(E|X - X_{0:i}|^2 / E|X|^2), \quad (2)$$

where  $X_{0:i}$  is the estimate of  $X$  reconstructed using only the first  $i$  components. A wavelet-based separable basis, with the spatial components  $\{S_j\}$  corresponding to Haar wavelets, is also included for comparison. We see that the error curves of the general and learned separable bases match almost exactly, indicating that the separable basis is equally efficient.

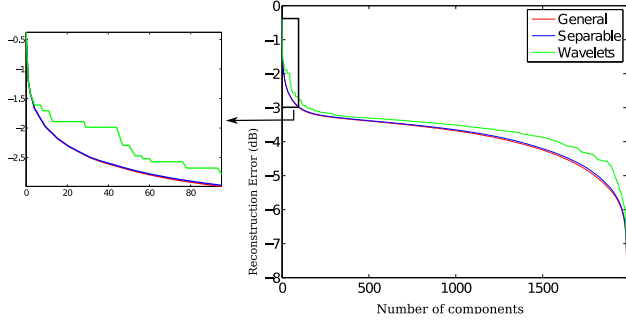


Figure 4. Comparison of different spatio-spectral basis sets in terms relative reconstruction error using a limited number of components. The figure compares a general basis set to one restricted to having separable spatial and spectral components. The separable basis has near identical reconstruction error to the general basis, indicating that it is equally efficient. A separable set with Harr wavelets as the spatial basis is also shown for comparison.

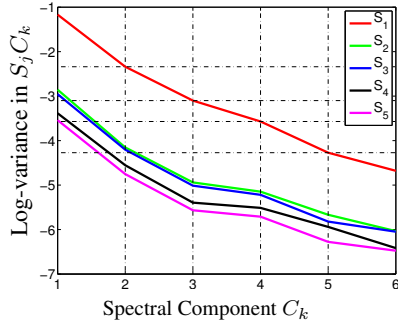


Figure 5. Variances in combinations with different spectral components, for the first five spatial components. The horizontal grid lines correspond to the values of the DC component  $S_1$ . Note that the different  $S_i$  have similar decays along the spectral dimension.

We next look at the variance along these separable components in Fig. 5, which compares the variance for the top spatial components when combined with each of the top spectral components. We note that the total variance in the different spatial components is distributed in similar proportions along the spectral dimension. Figure 6 provides another look at the variances of different components, and shows the relative ordering of the separable bases  $S_j[n]C_k[l]$  in terms of variance. Only the top fifteen separable components are shown for clarity of display. Note that combinations of the first spectral component with various spatial components have higher variance than the latter spectral components.

## 5. Coefficient Models

Having identified a separable spatio-spectral basis we now explore statistical models for coefficients in this basis. We look at distributions for each coefficient individually, as well as joint models for different spectral coefficients along the same spatial basis.

		Spatial Comp. $S_j$										
Spectral Comp. $C_k$		1	2	3	4	5	6	7	8	9	10	11
		1	3	4	6	7	9	10	11	12	13	15
2		2	14									
3		5										
4		8										

Figure 6. Top fifteen separable components  $V_i$  expressed in terms of combinations of the spatial and spectral components  $S_j$  and  $C_k$ . Note that combinations of the first spectral component  $C_1$  with various  $S_j$  often rank higher than the combinations of  $S_1$  (i.e. DC) with  $C_k, k > 1$ . This figure illustrates the relative importance of spatial to spectral resolution in terms of accurately representing a hyperspectral image.

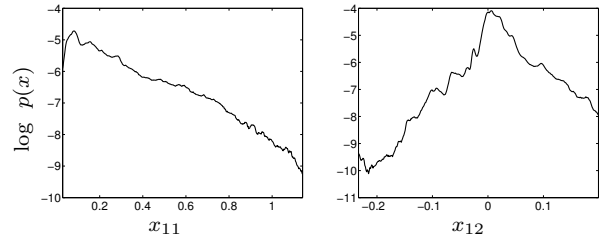


Figure 7. Empirical histograms for DC coefficients corresponding to different spectral components. In addition to having high variance, these coefficients show comparatively less structure than those corresponding to higher spatial components (see Fig. 8).

### 5.1. Modeling Individual Coefficients

Let  $x_{jk}$  be the coefficient of  $X$  in the basis component  $S_j[n]C_k[l]$ . We begin by looking at empirical distributions of the “DC” coefficients  $x_{1k}$  in Fig. 7. We find that these distributions differ qualitatively from those of the other coefficients (see Fig. 8), and exhibit comparatively less structure. In applications with grayscale and RGB images, DC (or “scaling”) coefficients are found to be poorly described by standard probability distributions and are often simply modeled as being uniform [6, 28], and the same could be done here.

The statistics of the higher spatial coefficients ( $x_{jk}$  for  $j > 1$ ) are more interesting. Figure 8 shows empirical distributions of  $x_{21}$  and  $x_{22}$  (the second spatial component). We see that these distributions are zero-mean, uni-modal, symmetric, and more kurtotic than a Gaussian distribution with the same variance, with heavier tails and a higher probability mass near zero. This matches intuition from grayscale and RGB image analysis that higher spatial sub-band coefficients are “sparse”.

We use a finite mixture of zero-mean Gaussians to model the distribution of these coefficients. Gaussian mixture models have been used for various applications with reasonable success [8, 28], and they have the advantage of al-



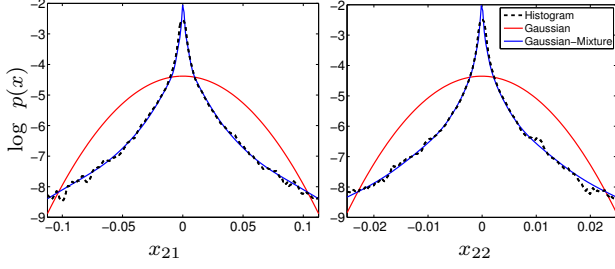


Figure 8. Distributions of different spectral coefficients corresponding to the second spatial component. The empirical distributions (shown in black) are uni-modal, symmetric, and heavier-tailed than a Gaussian distribution with the same variance (shown in red for comparison). They are well-modeled by a mixture of eight Gaussians (shown in blue).

lowing tractable inference. We define

$$p(x_{jk}) = \sum_{z=1}^Z p(z_{jk} = z) \mathcal{N}(x_{jk} | 0, \sigma_{jk,z}^2), \quad (3)$$

where  $z_{jk} \in \{1, \dots, Z\}$  is a latent index variable indicating that  $x_{jk}$  is distributed as a Gaussian with the corresponding variance  $\sigma_{jk,z}^2$ . Without loss of generality, we assume that the mixture components are sorted by increasing variance.

The model parameters  $\{p(z_{jk} = z), \sigma_{jk,z}^2\}_z$  are estimated from the database using Expectation Maximization (EM) [9], and in practice we find that a mixture of 8 Gaussians (*i.e.*,  $Z = 8$ ) provides accurate fits to the empirical distributions for all coefficients. These fits for the coefficients  $x_{21}$  and  $x_{22}$  are shown in Fig. 8.

## 5.2. Joint Models

Since the spatio-spectral basis vectors have been estimated through PCA, it follows that  $x_{jk}$  and  $x_{jk'}$  will be uncorrelated for  $k \neq k'$ , *i.e.*  $\mathbb{E}(x_{jk}x_{jk'}) = 0$ . However, given the model for individual coefficients in (3), this does not necessarily imply that they will be independent. Indeed, different spatial coefficients at the same spatial location in grayscale images are known to be related [8]. We now demonstrate that different spectral coefficients along the same spatial basis are also mutually dependent, and propose a model that encodes these dependencies.

We begin by examining whether knowing the value of the mixture index  $z_{jk}$  carries any information about the statistics of the coefficient  $x_{jk'}$  for a different spectral component  $k'$  along the same spatial basis  $j$ . We define  $\sigma_{jk'|z_{jk}}^2(z)$  to be the variance of  $x_{jk'}$  conditioned on the mixture index  $z_{jk}$  being  $z$ , and estimate it from a set of training patches  $\{X^i\}$  from the database, as

$$\sigma_{jk'|z_{jk}}^2(z) = \frac{\sum_i p(z_{jk} = z | x_{jk}^i) (x_{jk'}^i)^2}{\sum_i p(z_{jk} = z | x_{jk}^i)}, \quad (4)$$

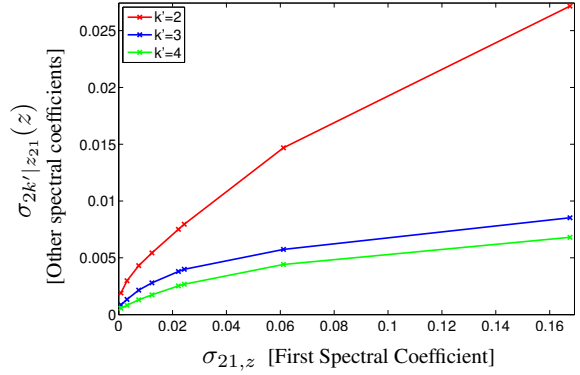


Figure 9. Relationship between the variances of different spectral coefficients  $\{x_{2k}\}$  for the same spatial basis  $S_2$ . We find that when  $x_{21}$  belongs to a mixture component having higher standard deviation  $\sigma_{21,z}$  (horizontal axis), the other spectral components  $x_{2k'}$  have higher standard deviations  $\sigma_{2k'|z_{21}}(z)$  (vertical axis) as well. This implies that the different spectral coefficients are not independent, because if they were, these curves would be horizontal.

where  $p(z_{jk} = z | x_{jk}^i)$  is computed for every training coefficient as

$$p(z_{jk} = z | x_{jk}^i) = \frac{p(z_{jk} = z) \mathcal{N}(x_{jk}^i | 0, \sigma_{jk,z}^2)}{\sum_{z'} p(z_{jk} = z') \mathcal{N}(x_{jk}^i | 0, \sigma_{jk,z'}^2)}. \quad (5)$$

Figure 9 shows these variances for different coefficients  $x_{2k'}$  conditioned on the mixture index  $z_{21}$  for the first spectral coefficient, and compares them to the corresponding mixture component variances  $\sigma_{21,z}^2$ . We see that the different spectral coefficients are indeed related. When the first spectral coefficient  $x_{21}$  belongs to a mixture component having higher variance, the expected variances of the other spectral coefficients  $\{x_{2k'}\}$  increase as well.

To capture this relationship, we update the model in (3) by including a joint distribution  $p(\{z_{jk}\}_k)$  on the mixture indices corresponding to different spectral coefficients along the same spatial basis as

$$p(\{x_{jk}\}_k) = \sum_{z_1, z_2, \dots} p(\{z_{jk} = z_k\}_k) \prod_k \mathcal{N}(x_{jk} | 0, \sigma_{jk,z_k}). \quad (6)$$

To fit this model, we first learn  $\{p(z_{jk} = z), \sigma_{jk,z}\}$  for each coefficient  $x_{jk}$  individually as before, and then we estimate the joint distribution of the indices  $p(\{z_{jk}\}_k)$  from the training patches  $\{X^i\}$  as

$$p(\{z_{jk} = z_k\}_k) \propto \sum_i \prod_k p(z_{jk} = z_k | x_{jk}^i). \quad (7)$$

Having fit this model, we can use the learned joint distribution of the mixture indices  $p(\{z_{jk}\}_k)$  to reason about the relationships between the corresponding coefficients  $\{x_{jk}\}_k$ . Figure 10 shows the estimated conditional distributions  $p(z_{2k'} | z_{2k})$  for different pairs of spectral coefficients

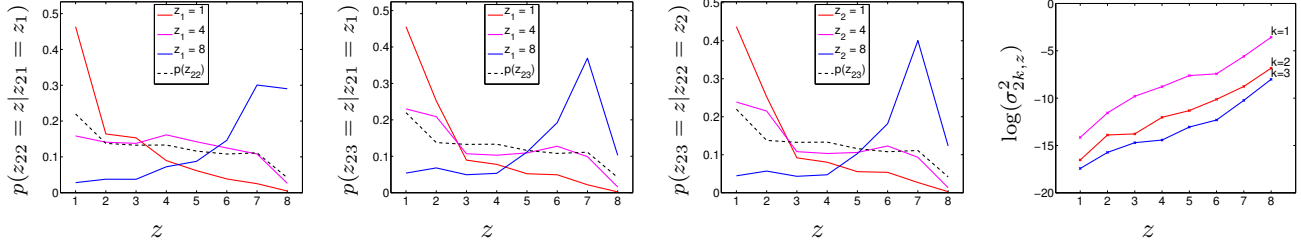


Figure 10. Conditional distributions of the mixture indices  $p(z_{2k'} | z_{2k})$  for different pairs of spectral coefficients along the same spatial basis  $S_2$ . The corresponding mixture component variances  $\sigma_{2k,z}^2$  are also shown for reference. Knowing the value of the mixture index  $z_{2k}$  for one spectral coefficient changes the distribution of the index  $z_{2k'}$ , corresponding to a different spectral coefficient, from the marginal distribution  $p(z_{2k'})$  (shown with dotted black line for comparison). Broadly, these graphs suggest that higher/lower magnitudes of one coefficient make higher/lower magnitudes respectively for other coefficients, along the same spectral basis, more likely.

along the spatial basis  $S_2$ . As expected, these distributions are different from the marginal distribution  $p(z_{2k'})$  (also shown for comparison). We find that conditioned on the mixture index  $z_{2k}$  having a value corresponding to higher mixture component variance, the index  $z_{2k'}$  for a different spectral coefficient  $x_{2k'}$  is more likely to correspond to higher variance mixture component as well, which is consistent with our observations in Fig. 9. Therefore, observing a high magnitude value for one coefficient makes a high value for another spectral coefficient along the same spatial basis more likely. This joint model can be exploited during inference, for example, when estimating a hyperspectral image from noisy or incomplete observations.

## 6. Discussion

In this work, we analyzed the joint spatial and spectral statistics of hyperspectral images using a new database of real-world scenes. We found that a separable basis, composed of independent spatial and spectral components, serves as an efficient representation for hyperspectral patches, and we studied the relative variance in these components. We then explored the statistical properties of coefficients in this basis and found that higher-frequency spatial components are accurately described by Gaussian mixture models. We also established that for the same spatial sub-band, different spectral coefficients are mutually dependent, and we described a joint distribution for mixture indices for different coefficients that encodes these dependencies.

A natural application of the statistical characterization described in this paper is in hyperspectral imaging. Acquisition systems should be constructed to exploit the interdependencies and correlations between different components so that they can efficiently acquire hyperspectral images with fewer measurements. General color filter array patterns (such as those proposed in [41]) can be designed to trade off spatial and spectral accuracy based on the relative variances of different components, and reconstruction methods can use the joint coefficient models during estimation. Similarly, these statistics are likely to be useful

when estimating “clean” hyperspectral images from observations degraded by noise, blur, chromatic aberration, *etc.* The database can be also used as “ground truth” to evaluate different strategies for these applications.

This paper presents a first look at spatio-spectral statistics and representations for hyperspectral images. Future work will include studying the statistics of specific classes of objects or regions in hyperspectral images, and leveraging these for vision applications. In addition to hyperspectral object models for recognition, understanding the difference in the statistics of homogenous regions with variations due to shading relative to that of regions that include material boundaries may be useful for segmentation and recovering “intrinsic images” [34].

Other avenues of future research include looking at representations derived using more sophisticated techniques such as independent component analysis and fields of experts [30], with the choice of representation likely to be geared towards specific vision tasks. We shall also explore using sparse codes, which have been previously proposed to describe spatial and spectral components independently in hyperspectral images [26]. Our observation about the mutual dependence between spectral coefficients for different spatial bands suggests that it would be useful to consider joint spatio-spectral coding strategies.

## Appendix: Camera-independent Statistics

As noted earlier, the analysis in the paper is performed relative to the camera’s sensitivity function shown in Fig. 1. Since this function is known, it is possible to compute the corresponding statistics for hyperspectral images captured by a different device with a different sensitivity, after making appropriate assumptions about the observation noise in the database. As a specific case of this, we look at properties of images captured by a hypothetical camera that has a flat sensitivity function. These can be interpreted as the properties of the underlying scene itself, without varying attenuation applied to the different wavelength bands.

Formally, we relate the captured hyperspectral patch

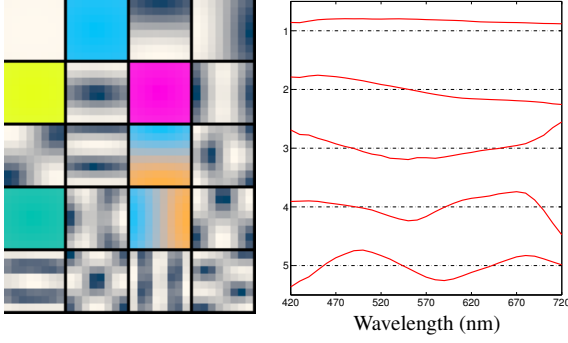


Figure 11. Basis vectors for hyperspectral patches from a camera with uniform sensitivity. Left: Most significant joint basis vectors, rendered in sRGB. Right: Spectral basis vectors  $\{C_{tk}\}$  that, combined with the spatial vectors  $\{S_j\}$  in Fig. 3, define an efficient separable basis.

$X[n, l]$  to the true un-attenuated version  $X_t[n, l]$  as

$$X[n, l] = s[l]X_t[n, l] + z[n, l], \quad (8)$$

where  $s[\cdot]$  is the known camera sensitivity, and  $z[\cdot]$  is observation noise. Following the analysis in Sec. 4, we seek to find the optimal basis for  $X_t$  through an eigen-decomposition of the covariance matrix  $\mathbb{E}X_tX_t^T$ . We assume white Gaussian noise,  $z[n, l] \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_z^2)$ , which gives us the following relation between the covariances of  $X$  and  $X_t$ :

$$\begin{aligned} \mathbb{E}X_t^2[n, l] &= s^{-2}[l] (\mathbb{E}X^2[n, l] - \sigma_z^2), \\ \mathbb{E}X_t[n, l]X_t[n', l'] &= s^{-1}[l]s^{-1}[l']\mathbb{E}X[n, l]X[n', l'], \\ &\text{if } n \neq n' \text{ or } l \neq l'. \end{aligned} \quad (9)$$

We use the values of  $\mathbb{E}XX^T$  estimated from our database, and we set the noise variance  $\sigma_z^2$  to be equal to half of its lowest eigen-value, *i.e.*, half the variance along the least significant basis vector. We can now compute the covariance matrix for  $X_t$ , and the optimal basis vectors thus obtained through PCA are shown in Fig. 11 (left).

Since  $X$  was shown in Sec. 4 to be represented efficiently using a separable basis and the camera sensitivity is the same for all pixels, it follows that the basis for  $X_t$  is also separable, and composed of the same spatial basis  $\{S_j[n]\}$  as for  $X$  and a spectral basis  $\{C_{tk}[l]\}$  shown in Fig. 11 (right). As expected, from comparing Fig. 11 to Fig. 3, we find that spectral basis vectors  $\{C_{tk}[l]\}$  for  $X_t$  represent an orthogonalized version of  $\{s^{-1}[l]C_k[l]\}$ .

Finally, we use our estimates of the covariance matrix of  $X_t$  to explore how efficient the human cone responses are at capturing the variance in the scenes in our database. We find that the sub-space spanned by the CIE XYZ vectors (designed to match the spectral response of human visual system) account for 77.22% of the total variance in  $X_t$ . In

comparison, the first three eigen-vectors  $\{C_{tk}\}_{k=0}^2$  account for 99.14% of the total variation. However, it is important to remember that human cone responses (as with any set of sensors) are restricted to have non-negative responses at all wavelengths. Also, the human visual system is likely to have evolved in different environments, and to be optimal for discriminative tasks that need not require capturing all the spectral variation.

## Acknowledgments

We wish to thank the reviewers for their useful suggestions, and Prof. David Brainard and Prof. Keigo Hirakawa for several insightful discussions. We would also like to acknowledge CRI Inc. for their technical support during this project, Brad Seiler for preliminary tests with the liquid crystal tunable filter, and Colleen Glenn and Siyu Wang for their assistance in collecting the database. Funding for this project was provided by the US Army Research Laboratory and the US Army Research Office under contract/grant number 54262-CI, as well as the National Science Foundation under Career award IIS-0546408.

## References

- [1] R. Baddeley, P. Hancock, and L. Smith. Principal components of natural images. *Network*, 3:61–70, 1992. 1
- [2] E. Belluco, M. Camuffo, S. Ferrari, L. Modenese, S. Silvestri, A. Marani, and M. Marani. Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. *Remote sensing of environment*, 105(1):54–67, 2006. 1
- [3] M. Borengasser, W. Hungate, and R. Watkins. *Hyperspectral remote sensing: principles and applications*. CRC, 2008. 1
- [4] J. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring from a single image using sparse approximation. In *Proc. CVPR*, 2009. 1
- [5] A. Castrodad, Z. Xing, J. Greer, E. Bosch, L. Carin, and G. Sapiro. Discriminative Sparse Representations in Hyperspectral Imagery. In *Proc. ICIP*, 2010. 1
- [6] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy beyond bags of pixels. In *Proc. CVPR*, 2008. 1, 4
- [7] H. Choi and R. Baraniuk. Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Trans. Imag. Proc.*, 10(9):1309–1321, 2002. 1
- [8] M. Crouse, R. Nowak, and R. Baraniuk. Wavelet-based statistical signal processing using hidden Markov models. *IEEE Trans. Sig. Proc.*, 46(4):886–902, 1998. 4, 5
- [9] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. of the Royal Stat. Soc. B*, 39(1):1–38, 1977. 5
- [10] D. Dicker, J. Lerner, P. Van Belle, S. Barth, D. Guerry, et al. Differentiation of normal skin and melanoma using high resolution hyperspectral imaging. *Cancer biology & therapy*, 5(8):1033, 2006. 1
- [11] H. Du, X. Tong, X. Cao, and S. Lin. A prism-based system for multispectral video acquisition. In *Proc. ICCV*, 2009. 1
- [12] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Imag. Proc.*, 15(12):3736–3745, 2006. 1

- [13] D. Foster, S. Nascimento, and K. Amano. Information limits on neural identification of colored surfaces in natural scenes. *Visual neuroscience*, 21(03):331–336, 2004. 1, 2
- [14] B. Gunturk, J. Glotzbach, Y. Altunbasak, R. Schafer, and R. Mersereau. Demosaicking: color filter array interpolation. *IEEE Sig. Proc. Magazine*, 22(1):44–54, 2005. 1
- [15] G. Heidemann. The principal components of natural images revisited. *IEEE Trans. PAMI*, 28(5):822–826, 2006. 1
- [16] K. Hirakawa and T. Parks. Adaptive homogeneity-directed demosaicking algorithm. In *Proc. ICIP*, 2003. 1
- [17] S. Hordley, G. Finalyson, and P. Morovic. A multi-spectral image database and its application to image rendering across illumination. In *Proc. Int. Conf. on Image and Graphics*, 2004. 1, 2
- [18] S. Lansel, M. Parmar, and B. A. Wandell. Dictionaries for sparse representation and recovery of reflectances. In *Proc. SPIE, Comp. Imaging VII*, 2009. 1, 2
- [19] T. Lee, T. Wachtler, and T. Sejnowski. The spectral independent components of natural scenes. In *Biologically Motivated Comp. Vis.*, 2000. 1, 2
- [20] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans. Imag. Proc.*, 17(1):53–69, 2007. 1
- [21] D. Marimont and B. Wandell. Linear models of surface and illuminant spectra. *JOSA A*, 9(11):1905–1913, 1992. 1, 2
- [22] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geoscience and Remote Sensing*, 42(8):1778–1790, 2004. 1
- [23] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio. Pedestrian detection using wavelet templates. In *Proc. CVPR*, 1997. 1
- [24] J. Park, M. Lee, M. Grossberg, and S. Nayar. Multispectral Imaging Using Multiplexed Illumination. In *Proc. ICCV*, 2007. 1
- [25] J. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of Munsell colors. *JOSA A*, 6(2):318–322, 1989. 1, 2
- [26] M. Parmar, S. Lansel, and B. A. Wandell. Spatio-spectral reconstruction of the multispectral datacube using sparse recovery. In *Proc. ICIP*, pages 473–476, Oct. 2008. 1, 2, 6
- [27] C. Párraga, T. Troscianko, and D. Tolhurst. Spatiochromatic properties of natural images and human vision. *Current Biology*, 12(6):483–487, 2002. 1
- [28] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Image denoising using Gaussian scale mixtures in the wavelet domain. *IEEE Trans. Imag. Proc.*, 12(11):1338–1351, 2003. 1, 4
- [29] L. Randeberg, I. Baarstad, T. Løke, P. Kaspersen, and L. Svaasand. Hyperspectral imaging of bruised skin. In *Proc. SPIE*, 2006. 1
- [30] S. Roth and M. Black. Fields of experts: A framework for learning image priors. In *Proc. CVPR*, 2005. 1, 6
- [31] R. Rowe, K. Nixon, and S. Corcoran. Multispectral fingerprint biometrics. In *Proc. Info. Assurance Workshop*, 2005. 1
- [32] D. Ruderman, T. Cronin, and C. Chiao. Statistics of cone responses to natural images: implications for visual coding. *JOSA A*, 15:2036–2045, 1998. 1
- [33] G. Stamatas, C. Balas, and N. Kollias. Hyperspectral image acquisition and analysis of skin. In *Proc. SPIE*, 2003. 1
- [34] M. Tappen, W. Freeman, and E. Adelson. Recovering intrinsic images from a single image. *IEEE Trans. PAMI*, pages 1459–1472, 2005. 6
- [35] E. Underwood, S. Ustin, and D. DiPietro. Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment*, 86(2):150–161, 2003. 1
- [36] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. Imag. Proc.*, 4(11):1549–1560, 2002. 1
- [37] J. Van de Weijer and T. Gevers. Color constancy based on the grey-edge hypothesis. In *Proc. CVPR*, 2005. 1
- [38] T. Wachtler, T. Lee, and T. Sejnowski. Chromatic structure of natural scenes. *JOSA A*, 18(1):65–77, 2001. 1
- [39] Y. Weiss and W. Freeman. What makes a good model of natural images? In *Proc. CVPR*, 2007. 1
- [40] G. Wyzecki and W. Stiles. Color Science. *Concepts and Methods, Quantitative Data and Formulae*, 1982. 2
- [41] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar. Generalized assorted pixel camera: Post-capture control of resolution, dynamic range and spectrum. Technical Report CUCS-061-08, Columbia University, 2008. 1, 2, 6