SPATIO-SPECTRAL RECONSTRUCTION OF THE MULTISPECTRAL DATACUBE USING SPARSE RECOVERY

Manu Parmar and Steven Lansel

Electrical Engineering Department,
Stanford University,
Stanford, CA-94305, USA.

Brian A. Wandell

Psychology Department,
Stanford University,
Stanford, CA-94305, USA.

ABSTRACT

Multispectral scene information is useful for radiometric graphics, material identification and imaging systems simulation. The multispectral scene can be described as a datacube, which is a 3D representation of energy at multiple wavelength samples at each scene spatial location. Typically, multispectral scene data are acquired using costly methods that either employ tunable filters or light sources to capture multiple narrow-bands of the spectrum at each spatial point. In this paper, we present new computational methods that estimate the datacube from measurements with a conventional digital camera. Existing methods reconstruct spectra at single locations independently of their neighbors. In contrast, we present a method that jointly recovers the spatio-spectral datacube by exploiting the data sparsity in a transform representation.

Index Terms— Multispectral imaging, sparse recovery

1. INTRODUCTION

A full description of the 3D multispectral scene datacube is useful for many applications like scene rendering under arbitrary illuminants, colorimetric analysis, machine vision tasks, etc. Another important application of spectral scene information is in the design and evaluation of imaging pipeline components. Simulation tools (e.g., [1]) can use multispectral source data to analyze the impact on final image quality of factors like chromatic aberration in the lens, the response of the imaging sensor, etc.

The datacube is the 3D signal corresponding to a scene with m rows, n columns, and k bands. Each voxel of the scene is addressed by its 2D spatial coordinates and a third wavelength coordinate. The wavelength dimension represents the energy distribution along the visible range of the electromagnetic spectrum ($\sim [400, 700]$ nm). Typical image acquisition systems sample the spectrum at each spatial location using a limited number of sampling functions; typically, the RGB channels. A significant amount of scene spectral information is lost in the process. Figure 1 shows an illustration of a multispectral scene and the corresponding 3-color RGB image.

Fig. 1. An illustration of the multispectral datacube with ten bands and the corresponding image as it would be acquired with a color imager.

The most accurate class of multispectral imaging methods found in the literature acquire multispectral scenes by using liquid crystal tunable (LCT) filters [2] and a monochromatic camera (usually of scientific-grade). The datacube is recovered from acquisitions of several narrow spectral bands of the scene. At each spatial location, the spectrum is estimated from these data using linear reconstruction techniques. Such LCT based multispectral imaging systems are costly both in terms of equipment costs and the training required to use them. Methods have also been proposed that use commonly available optical filters and 3-color digital still cameras (e.g. [3]) to reconstruct the datacube. Typically, 3-color cameras use a color filter array (CFA) to spatially subsample the signal and each acquisition with the camera gives one of 3 distinct samples of the spectrum at each spatial location. Such methods address the datacube reconstruction problem by independently reconstructing the spectrum at each spatial location after either demosaicing or downsampling CFA data for each acquisition.

In this paper we present a novel method for the reconstruction of the datacube. We first present a method to reconstruct spectra by ℓ_1-norm minimization in a sparse representation. Using these ideas, we develop a sparse representation of the spatio-spectral datacube and propose a method for sparse recovery of the datacube. Finally, we implement this method to reconstruct the datacube from data acquired using multiple acquisitions from a commonly available digital camera.

2. SPECTRAL RECONSTRUCTION

We consider first the problem of reconstructing the spectrum from measurements made with a limited number of sampling functions. Let $z \in \mathbb{R}^k$ represent the discretized spectrum at

a particular location. A number of measurements of z may be acquired by modulating the response of the imaging sensor by using optical filters or illuminants with different SPDs. Let $c \in \mathbb{R}^t$ denote the measurements acquired at a particular spatial location. Then,

$$c = F^T z + \nu, \qquad (1)$$

where the columns of $F \in \mathbb{R}^{k \times t}$ are the combined sensitivities of the imaging sensor and optical filters (or illuminants), ν is the system noise. The spectrum at a particular point can be recovered from color measurements by solving the problem

$$\hat{z} = \arg\min_z \phi(c, z), \qquad (2)$$

where ϕ is some objective function that describes the error between actual and estimated values of z. Typically, $t \ll k$, and the problem is ill-posed.

3. SPARSE RECOVERY OF SPECTRA

Recently, a number of applications in signal processing, e.g., denoising, signal recovery, MRI, etc. have been proposed based on the ideas of compressed sensing [4]. Many common signals have the property that they can be represented with only a few non-zero coefficients in some transform basis with little loss of information. In the context of object spectra, suppose a transform basis W exists where object spectra are sparsely represented. This implies that there exists a representation a of a spectrum $z \in \mathbb{R}^k$, such that $||z - Wa||_2 < \epsilon$, where only p elements of a are non-zero, and $p \ll k$. An ℓ_1-regularized solution to the spectrum estimation problem may then be found as $\hat{z} = W\hat{a}$, where

$$\hat{a} = \arg\min_a ||c - F^T W a||_2 + \tau ||a||_1. \qquad (3)$$

The ℓ_1-norm of a is the regularizer and τ is the a parameter that weights sparsity in relation to data fidelity.

3.1. Sparse representations of spectra

For the spectral estimation problem, it becomes necessary to find a suitable sparsity basis for z. Note that ℓ_1-regularization can be also be successfully applied when W is not an orthonormal basis, and is overcomplete $\left(W \in \mathbb{R}^{k \times q}, q > k \right)$ [5]. In such a case, W is no longer a strict basis, but is referred to as a *dictionary* and its columns are known as *atoms*. Often, such redundant dictionaries will lead to sparse representations useful for solving inverse problems.

Object (and most illuminant) spectra, as a class of signals, are inherently lower-dimensional. To put this in perspective, reflectance is typically considered to be well sampled when sampled 31 times in the interval [400,700] nm. It has been shown that most spectra may be represented by coefficients corresponding to as few as 6 basis vectors [6]. An appropriate compact representation is easily obtained by truncating the right eigenbases found using the SVD of a set of representative spectra. Equivalently, if spectral statistics are known, a compact representation is found by truncating the corresponding basis found using PCA. This same PCA basis may be used to generate a sparse representation by considering the entire basis and constraining the ℓ_0-norm (the number of nonzero elements) of
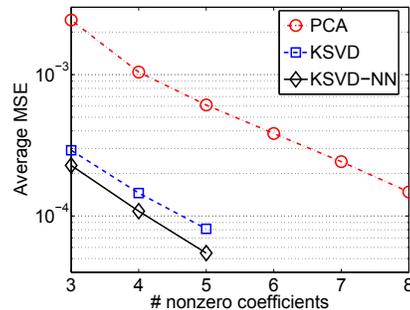

Fig. 2. Average MSE for a test set of 100 spectra. For equal MSE, the KSVD-NN dictionary represents the spectra with fewer coefficients than the PCA basis.

a to some small value. Dictionaries may be learned empirically from representative data by solving other mathematical problems. The K-SVD algorithm [7] may be used to derive dictionaries for sparse representations of signals by solving an ℓ_2-minimization similar to the k-means clustering problem. The solution is subject to a constraint on the ℓ_0-norm of the coefficients and results in sparse decompositions.

We have investigated a number of sparse representations for object spectra. Although we do not describe here in detail the process of choosing an appropriate basis for sparse recovery of spectra, a more complete treatment can be found in [8]. Here, we summarize the performance in terms of representation errors of 3 different sparse representations – the PCA basis, the KSVD learned dictionary, and the KSVD dictionary with nonnegativity constraints (KSVD-NN). In each case the dictionary was derived from training a set of 450 measured reflectance spectra of common objects. These spectra were measured with a spectraradiometer and contain samples of common color test charts, skin, and other natural objects [3] that were sampled every 10 nm in the interval [400,700] nm to give length 31 vectors. We assume that this set of spectra is representative of all spectra we wish to recover from undersampled measurements.

Figure 2 shows the average error for representations with different levels of sparsity found for a set of 100 test spectra distinct from the training set. The average representation error for a particular level of sparsity is found over the test set by constraining the number of nonzero coefficients in $\hat{a}$. At equivalent MSE levels, both the KSVD and KSVD-NN dictionaries well-represent the test spectra with far fewer nonzero coefficients than the PCA basis. For instance, Both KSVD-based dictionaries have an average MSE with 4 nonzero coefficients that is similar to the MSE with the PCA basis with 8 nonzero coefficients. Based on these experiments, the KSVD-NN dictionary was chosen for sparse recovery.

3.2. Simulations

Simulations were carried out for evaluating the sparse recovery scheme. A number of different measurement systems were designed by using different numbers of spectral filters that are equally spaced Gaussian curves along the wavelength dimension. For each such system, measurements were simulated according to the forward model in (1) with a number of different levels of additive white Gaussian noise. A measurement scheme is then defined by the number of filters and the level of noise.
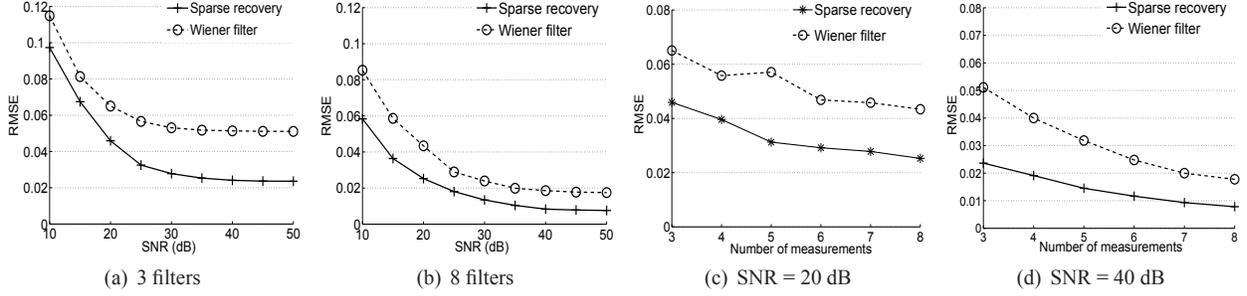
(a) 3 filters   (b) 8 filters   (c) SNR = 20 dB   (d) SNR = 40 dB

Fig. 3. The effect of SNR and number of wavelength samples on spectrum reconstruction error. Average RMSE over the test set of spectra as a function of SNR for measurements with 3 (panel a) and 8 (panel b) wavelength sampling functions. Average RMSE over the test set as a function of wavelength samples at 20 dB (panel c) and 40 dB (panel d).

For each measurement scheme, we sampled and reconstructed the 100 spectra in our test set using 2 methods – sparse recovery with the KSVD-NN dictionary by solving (3), and the Wiener filter (Shimano [9] has recently proposed a Wiener filter-based method to recover spectra without prior knowledge of scene or noise statistics.). We used the Gradient Projection for Sparse Reconstruction (GPSR) algorithm [10] to perform the minimization in (3). The Wiener filter solution is found as

$$\hat{z} = R_z F \left(F^T R_z F + R_\nu \right)^{-1} c, \qquad (4)$$

where the autocorrelation matrix R_z is found from our training data of 450 spectra and it is assumed that $R_\nu = \sigma_\nu^2 I$, where σ_ν^2 is the noise variance and I is the identity matrix.

Figures 3 shows the RMSE obtained over the test set for different SNR levels for measurements carried out with 3 (Fig. 3(a)) and 8 (Fig. 3(a)) evenly spaced Gaussian filters respectively. Note that in each case sparse recovery with noisy measurements performs comparably with the Wiener filter with high SNR measurements. The RMSE levels using different number of filters for SNR levels 20 dB (Fig. 3(c)) and 40 dB (Fig. 3(d)) are also shown.
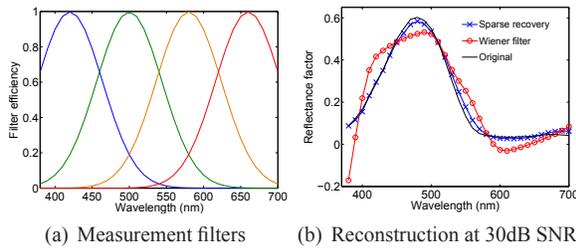

(a) Measurement filters   (b) Reconstruction at 30dB SNR

Fig. 4. (a) The filters that form the columns of F, and (b) reconstruction results for a particular spectrum with the proposed sparse recovery method and the Wiener filter. The RMSE for sparse recovery is near the median of the sample population.

The performance of sparse recovery is illustrated in Fig. 4 for a particular measurement system and a spectrum. The measurements scheme uses only 4 equally spaced Gaussian-shaped spectral sampling functions (Fig. 4(a)) and an SNR of 30 dB. For this spectrum, the RMSE for the sparse recovery case is 0.018 (45th percentile in the test set) while the RMSE for the Wiener filter is 0.082.

4. SPATIO-SPECTRAL RECOVERY

In Section 3 we established that the sparsity property can be used effectively for the recovery of spectra from noisy, undersampled measurements. Consider now the spatial dimensions: natural images, as a class of 2D signals, are inherently sparse or compressible in various representations (Fourier, wavelet, etc.). In the context of the scene datacube, each of the k bands of the cube may be considered to be a 2D signal akin to a grayscale image. Each band may then be transformed into a wavelet basis and be represented compactly. It follows from the compressibility in both the spatial and spectral dimensions considered individually, that the datacube is compressible in some appropriate sparsity basis when considered jointly in 3D. In this section we propose such a basis and a method to recover the datacube from subsampled measurements by exploiting the joint spatio-spectral 3D compressibility.

For notational convenience we represent the data cube as the vector $d \in \mathbb{R}^{mnk}$ formed by stacking the column-ordered bands of the 3D spectral data. Consider now Ψ, the sparsity basis of d such that $d = \Psi x$; x are the sparse coefficients. We use the 3D sparse representation derived from the KSVD-NN dictionary from Section 3 along the spectral dimension and the 2D Haar wavelet basis along the spatial dimensions. The Haar wavelet basis is chosen because of its conceptual and computational simplicity. The measurement system is defined by the matrix Φ, which incorporates both spatial and spectral sampling.

The sparse signal x is estimated as the sparsest signal that agrees with the data. The problem may be formulated as

$$y = Ax + \eta, \qquad (5)$$

where $A = \Phi^T \Psi$ and η is measurement noise. We define an objective function as

$$J(x) = ||y - Ax||_2^2 + \tau ||x||_1, \qquad (6)$$

such that an estimate of the sparse signal x may be obtained as

$$\hat{x} = \arg \min_x J(x). \qquad (7)$$

The parameter τ places a weight on the penalty on the density of x. The datacube is recovered as $\hat{d} = \Psi \hat{x}$.
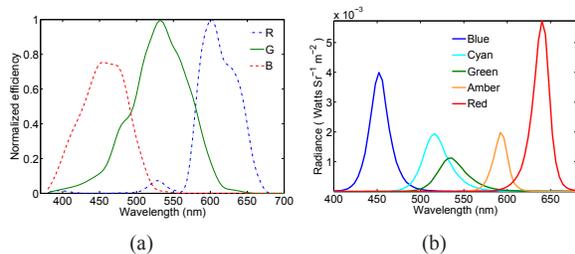
Fig. 5. (a) Normalized spectral efficiency functions of the camera channels, (b) the radiances of the illuminants measured at the center of the scene.

Fig. 6. (a) Multispectral scene rendered in RGB, and (b) Comparison of measured and recovered reflectances of the MCC *yellow* and *purplish blue* patches.

5. EXPERIMENTS

We used a Nikon D2Xs digital SLR camera for our experiments. The Nikon D2Xs has a 12-bit ADC and makes available camera RAW images that have not been processed for color-balancing, gamma, etc. The camera uses an RGGB Bayer-pattern color filter array [11]. The spectral transmittances of the color filters of this camera were experimentally determined and are shown in Fig. 5(a). One acquisition with this camera gives 3 spatially subsampled 2D images that represent measurements due to the filters shown in Fig. 5(a). To improve the spectral recovery capability of the system we take multiple acquisitions with scenes that are illuminated with an LED-based lighting system that has 5 different illuminants in the visible range [12]. The SPDs of these LEDs as measured at the center of the imaged scene are shown in Fig. 5(b). With this arrangement, ideally, we would have 5 spectral measurements (camera color sensitivity modulated by the LED SPDs) at each spatial location for a total of 15 distinct filter measurements. In practice, only 8 of these channels give useful signals; the rest have very low SNRs.

The datacube measurement system Φ now incorporates the spatial sampling due to the Bayer pattern and the 8 spectral filters. The sparsity basis Ψ incorporates the 2D Haar wavelet basis applied to the spatial dimensions and the KSVD-NN dictionary along the spectral dimension. The GPSR algorithm [10] is specifically designed to solve the optimization problem in (7). We used the monotonic version of the GPSR-BB algorithm to perform the minimization of $J(x)$.

Figure 6(a) shows an RGB representation of the scene datacube obtained by using the proposed method. A Macbeth Color Checker$^©$ (MCC) test chart was included in the scene to enable comparison of recovered spectra with known reflectances. Figure 6(b) shows this comparison for the *yellow* and *purplish blue* patches of the MCC.

6. CONCLUSIONS

In this paper we have presented a novel method to estimate object spectra from undersampled, noisy measurements using sparse recovery techniques. This approach is extended in a method to recover the entire 3D multispectral scene datacube from only a few acquisitions taken with a conventional digital still camera. The proposed method exploits the inherent compressibility of the data cube in both the spectral and spatial domains. We show the results of our method on experimental data acquired with a consumer-grade digital camera.
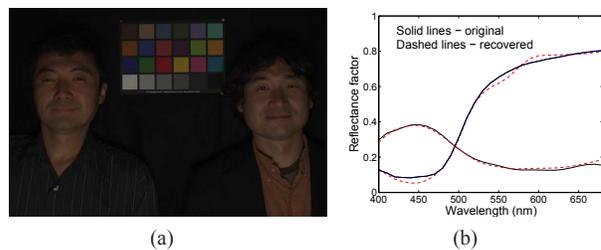
7. REFERENCES

[1] J. E. Farrell *et al.*, "A simulation tool for evaluating digital camera image quality," in *Image Quality and System Performance, Proceedings of the the SPIE*, vol. 5294, Jan. 2004, pp. 124–131.

[2] J. Y. Hardeberg *et al.*, "Multispectral image capture using a tunable filter," *Optical Engineering*, vol. 41, no. 10, pp. 2532–2548, Oct. 2002.

[3] M. Parmar *et al.*, "A database of high dynamic range visible and near-infrared multispectral images," in *Digital Photography IV, Proceedings of the SPIE*, vol. 6817, Mar. 2008, pp. 68 170N – 68 170N–10.

[4] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[5] D. L. Donoho *et al.*, "Stable recovery of sparse overcomplete representations in the presence of noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 1, pp. 6–18, 2006.

[6] D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra," *J. Opt. Soc. Am. A*, vol. 9, no. 11, pp. 1905–1913, 1992.

[7] M. Aharon *et al.*, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[8] S. Lansel *et al.*, "Dictionaries for sparse representation and recovery of object spectra," in *Computational Imaging VII, IS&T/SPIE conference on Electronic Imaging (submitted)*, 2009.

[9] N. Shimano, "Recovery of spectral reflectances of objects being imaged without prior knowledge," *Image Processing, IEEE Transactions on*, vol. 15, no. 7, pp. 1848–1856, Jul. 2006.

[10] M. A. Figueiredo *et al.*, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[11] B. Bayer, "Color imaging array," U.S. Patent 3971065, July 1976.

[12] M. Parmar *et al.*, "An LED-based multispectral imaging system," in *Digital Photography V, IS&T/SPIE conference on Electronic Imaging (submitted)*, 2009.