THE VALUE OF MULTISPECTRAL OBSERVATIONS IN PHOTON-LIMITED QUANTITATIVE TISSUE ANALYSIS

Zachary T. Harmany, Xin Jiang, and Rebecca Willett

Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA

ABSTRACT

Multispectral fluorescence data can be used within sparse decomposition methods to separate key cellular structures, even when the number of photons per spectral band is very small. However, there are two key costs associated with multispectral data acquisition: (a) for a fixed data acquisition time, increasing the number of spectral bands means decreasing the number of photons (and hence SNR) per band, and (b) the optical system becomes more complex and expensive. These costs lead to important tradeoffs between the information content and the noise of the observations. This paper describes a mathematical framework for assessing this tradeoff and supporting experimental results.

Index Terms— Photon-limited imaging, Poisson noise, sparse decomposition, microscopy, fluorescence imaging, convex optimization, spectral imaging

1. INTRODUCTION

"Margin assessment" is a critical problem when physicians are excising tumors from cancer patients. After removing a tumor, the medical team must accurately determine whether all the tumor tissue has been removed. The alternative of completing the surgery, conducting non-invasive screening post-operation, and then scheduling follow-up surgery is both risky to the patient and expensive.

To aid with margin assessment *during surgery*, engineers have developed fluorescent micro-endoscopes which can be used to probe and image tissue around a small incision quickly. However, due to fluorescence and tight time constraints, the observations are often photon-limited. From this data, we wish to estimate a decomposition of the image into nuclei, muscle, and fat (adipose) components. Computational tissue analysis along these lines has been considered previously for relatively low noise settings using sparse decomposition techniques [1]. In our context, photon limitations are a dominant source of errors. However, we have limited control of the diversity of our observations via the ability to collect multispectral observations. This paper explores the tradeoffs between photon limitations and the measurement diversity associated with multispectral observations.

In particular, assume that the fluorescence occurs in the range of wavelengths $0 \leq \lambda \leq \Lambda$, and that the spectral signatures of the different tissues and fluorescent dyes are known (for example, based on prior laboratory measurements from a spectrometer). We must then decide on the ideal spectral resolution of the instrument – *i.e.,* the number of bins of wavelengths, M, spanning $[0, \Lambda]$; this decision introduces a key tradeoff between measurement diversity and the signal-to-noise ratio. Specifically, larger M means that each spectral bin will contain fewer observed photons and will hence be noisier, but it also provides us with a more disparate set of measurements which can aid in separating different tissue components. This leads to the following question addressed by this paper: *What is the information value associated with multi-spectral observations for photon-limited image decomposition?*

The work described in this paper has several interesting connections to the authors' previous work on Poisson compressed sensing [2]. In that problem, we saw that there was a critical number of measurements required for accurate reconstruction of the underlying sparse signal, and that as we increased the number of measurements beyond this critical threshold, the reconstruction quality decreased along with the SNR of each observation. The authors developed a theoretical analysis of this phenomenon. This paper empirically demonstrates a similar effect in a real-world Poisson inverse problem.

2. PROBLEM FORMULATION

Here we describe a general model for multi-wavelength imaging. Each $x_i \in \mathbb{R}^n$ is a vector representation of an n-pixel image. In this model, we have N tissue components $\{x_j\}_{j=1}^N$ that combine linearly to produce M images $\{z_i\}_{i=1}^M$ acquired using different wavelengths or spectral bands:

$$z_i = \sum_{j=1}^{N} A_{i,j} x_j, \; i = 1, \ldots, M, \tag{1}$$

where $A \in \mathbb{R}^{M \times N}$ is the multispectral mixing weight matrix. Element $A_{i,j}$ describes how brightly tissue component j appears in spectral band i. More specifically, assume tissue type j has spectrum $f_j(\lambda)$, and let

$$A_{i,j} \triangleq \int_{(i-1)\Lambda/M}^{i\Lambda/M} f_j(\lambda) d\lambda.$$

Then the j^{th} column of A, denoted a_j, corresponds to the spectrum of tissue type j discretized to M elements. In this way we can write $A = [a_1, \cdots, a_N]$. If we let $x = (x_1, \ldots, x_N)$ and $z = (z_1, \ldots, z_M)$, then this model can be more succinctly expressed as

$$z = (A \otimes I_n)x, \tag{2}$$

where $\otimes$ denotes the Kronecker matrix product and I_n is the size-n identity matrix. We then acquire photon-limited observations of the M images, which can be accurately described by a inhomogeneous Poisson process [3]:

$$y \sim \text{Poisson}(z). \tag{3}$$

In order to facilitate the recovery of the tissue components x from the data y, we utilize the assumption that each tissue component x_i

This work was supported by NSF CAREER Award No. CCF-06-43947, NGA Grant HM1582-10-1-0002 and NSF Grant DMS-08-11062.

(a) True Nuclei (b) True Muscle (c) True Haze

Fig. 1. Tissue components

can be sparsely represented or approximated in a dictionary $D_i \in \mathbb{R}^{n \times p_i}$, meaning that

$$x_i = D_i \theta_i, \; i = 1, \ldots, N \tag{4}$$

where the coefficient sequence θ_i is mostly comprised of very few large coefficients. We allow for the possibility that each dictionary may be over-complete ($p_i > n$) or under-complete ($p_i < n$). Again we may compactly express this as $x = D\theta$, where $D = \text{bdiag}(D_1, \ldots, D_N)$ is a block-diagonal dictionary matrix and $\theta = (\theta_1, \ldots \theta_N) \in \mathbb{R}^p$ ($p \equiv \sum_{i=1}^N p_i$) is the concatenation of the coefficients. In this paper, we assume that each dictionary is a 1-tight (Parseval) frame, although extensions to other cases are possible. This assumption ensures that $D_i D_i^T = I_n$ for all $i = 1, \ldots, N$, as is true for curvelet dictionaries, and (normalized) concatenations of orthonormal bases. From this, we will also have that $DD^T = I_{nN}$.

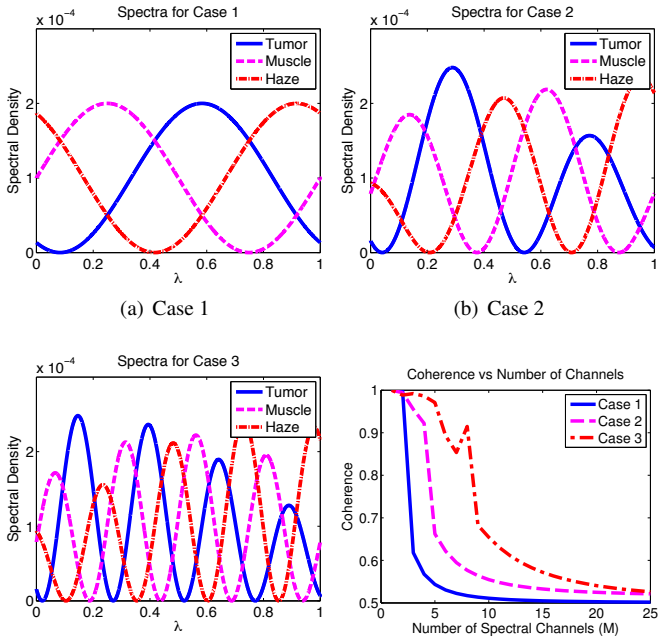
With this model and a set of assumptions, we recover x by solving the sparsity-regularized maximum likelihood problem

$$\begin{aligned} \underset{\theta}{\text{minimize}} \quad & -\log p(y|(A \otimes I_n)D\theta) + \|\tau \odot \theta\|_1 \\ \text{subject to} \quad & D\theta \geq 0, \end{aligned} \tag{5}$$

where $\tau \in \mathbb{R}_+^p$ is a vector of regularization parameters, and $\odot$ denotes componentwise (Hadamard) multiplication. The common case (and used in our numerical experiments) is that there is a single regularization parameter associated with each $\theta_i, i = 1, \ldots, N$, which controls the relative sparsity tradeoff in each dictionary. In this simplified case, the sparsity penalty may be written as $\sum_{i=1}^N \tau_i \|\theta_i\|_1$. The nonnegativity constraints are required since each of the tissue component florescence images are Poisson intensities, which are required to be nonnegative. We briefly describe an approach that can be used to effectively solve this minimization in Sec. 4.

3. TRADEOFFS

Let $R \triangleq (A \otimes I_n)D$ denote the full spatio-spectral sensing matrix. As we show below, depending on the dictionaries and spectral profiles available, the number of spectral bands (M) necessary to ensure reliable reconstruction of θ from y depends critically on the coherence properties of this resulting sensing matrix. For example, if $R^T R = I_{nN}$, then in the absence of noise, the different tissue components can be perfectly separated from the observations. More generally, we need to quantify how close $R^T R$ is to I_{nN}, and how the number of spectral bands influences this distance. The analysis below shows that the structure of $R^T R$ depends on both the structure of the sparsifying dictionaries and the weight matrix A. Usually increasing M results in better coherence properties. In a practical setup, however, this improvement is mitigated by the fixed amount of total available light and its dispersion across the M spectral bands.



(a) Case 1 (b) Case 2

(c) Case 3 (d) Coherence vs. M

Fig. 2. Spectral profiles of three simulated tissue components in three cases (a)-(c), and (d) coherence behavior as a function of the number of spectral bands, M, for each case.

3.1. Incoherence of multispectral measurements of tissue mixtures

We define the Gram matrix as $G \triangleq R^T R$. When G is close to an identity matrix, we can expect a better performance in reconstructing θ. This is formally expressed using notions of *incoherence* [4, 5] or the *Restricted Isometry Property* [6] in the compressed sensing literature. Here we define the *coherence* of a matrix B as

$$\mu(B) = \max_{i \neq j} \frac{b_i^T b_j}{\|b_i\| \|b_j\|} \tag{6}$$

where b_i denotes the ith column of B. A sensing matrix that has low coherence $\mu(R)$ translates to a Gram matrix that is close to an identity matrix.

In our setting, Kronecker product properties can be used to show

$$G = R^T R = [(A \otimes I_n)D]^T [(A \otimes I_n)D] = D^T [(A^T A) \otimes I_n] D,$$

which can be partitioned into blocks of the form $a_i^T a_j D_i^T D_j$. This shows that the closer each block is to zero for $i \neq j$, the closer G is to an identity matrix. When the sampled spectra of the tissue components are very similar, we may only resolve these tissue types if there exist correspondingly incoherent sparsifying bases for each tissue type. Consider the tissue components in Fig. 1. Even if the spectral characteristics are similar, it may be possible to resolve isolated nuclei (sparse in the spatial domain) against a background of longitudinal muscle fibers (sparse in the DCT domain) simply due to their incoherent dictionaries. The additional flexibility of the spectral response can be leveraged when two tissue types are sparse in very similar bases, such as muscle tissues, adipose (fat) tissue, and the haze component (described in more detail in the simulations section).

This analysis yields the property $\mu(R) \leq \mu(A)\mu(D)$. This allows us to simply analyze the coherence of the multispectral mixing matrix as a worst-case analysis for our simulations section. In particular, we examine the coherence behavior versus M using three different spectral signatures for the tissue types. Each of the three sets of spectra are shown in Fig. 2. In each of the different cases, the spectra oscillate more rapidly, which translates to requiring more spectral channels to separate out the different spectral profiles. For instance, if we select a target coherence level of 0.55, we need $M = 5$, 12, and 20 spectral channels to achieve this level for Case 1, 2, and 3 respectively.

3.2. Connections to Poisson compressed sensing

In previous work [2], we examined the performance of compressed sensing in the presence of photon noise, the dominating noise measure in a spectral imaging system with a large number of spectral bands. We demonstrated that if R satisfies a variant of the restricted isometry property, so that $R^T R \approx I$, then a reconstruction of the form

$$\widehat{\theta} = \arg\min_{\theta} - \log p(y|R\theta) + \tau\|\theta\|_1$$

resulted in a mean squared error

$$\mathbb{E}[\|\widehat{\theta} - \theta^*\|_2^2] \preceq Mn \left(\frac{\log Nn}{T}\right)^{2\alpha/(2\alpha+1)} + \frac{\log(N/M)}{Mn}, \quad (7)$$

where τ is a tuning parameter, $T \triangleq \sum_i (D\theta)_i$ is the total intensity of the scene (proportional to the total amount of time required to collect the photon data), and $\alpha > 1$ reflects the sparsity or compressibility of the θ vectors. The bound in (7) suggests that M should be chosen as small as possible (while still ensuring the necessary degree of incoherence) to ensure that the estimate has as little error as possible. This is because larger M causes the limited number of photons to be dispersed across more detectors and limits the signal-to-noise ratio of the observations.

4. ALGORITHMS

In the Poisson observation model, estimating θ from y using the minimization in (5) can be a significant computational challenge. The authors developed methods for solving this inverse problem, where (a) the number of unknowns may potentially be larger than the number of observations and (b) x admits a sparse approximation or representation [7]. Our optimization formulation uses a penalized negative Poisson log-likelihood objective function with nonnegativity constraints (since Poisson intensities are naturally nonnegative). In particular, our approach incorporates key ideas of using separable quadratic approximations to the objective function at each iteration and penalization terms related to ℓ_1 norms of coefficient vectors, total variation seminorms, and partition-based multi-scale estimation methods. Our approach is very stable relative to alternative methods based on augmented Lagrangians (AL) because there is no need to choose a good value of the AL parameter, which is a significant practical challenge [8].

5. SIMULATIONS

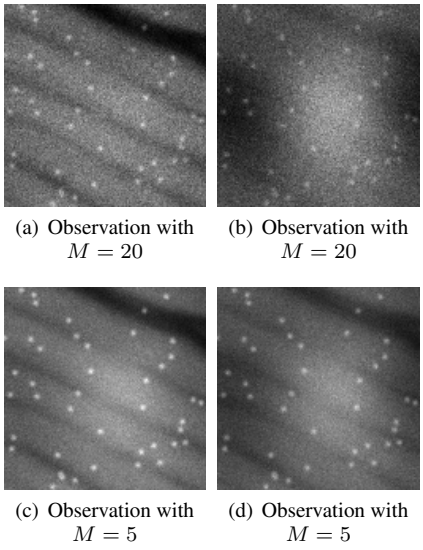
5.1. Simulation Setup

In our simulations, we consider two basic tissue components: nuclei and muscle fibers. In pathology, these nuclei may correspond to healthy normal tissue, or possibly a malignant tumor. Our goal is to separate out these nucleus locations for further processing (e.g., determination of their size and density can be diagnostically relevant). In addition, we consider a component that corresponds to a background of non-binding contrast agent we refer to as *haze* due to visual artifacts it causes in the images.

The simulated nuclei image is generated by passing some randomly located bright dots through a round low-pass Gaussian filter. The muscle component is taken from a real muscle tissue image followed by minor processing in order to eliminate unwanted vessel and fiber components while preserving natural muscle texture. As the haze refers to the general blurred background in the observations, we use low-pass filtered Gaussian noise to simulate this component. The generated components are shown in Fig. 1 (a)–(c).

The fact that different components have different spectral densities (brightness) gives us the chance to achieve better performance via adding additional spectral bands. We consider three different sets of spectra in our simulations, and are shown in Fig. 2 (a)–(c). Fig. 2 (d) shows the coherence $\mu(A)$ of sensing matrix A versus the number of spectral bands, where we defined coherence in (6). Generally, the more spectral bands we have, the lower the coherence level. As addressed in Sec. 3, lower coherence benefits reconstruction performance.

We constrain the total observation time to be constant. That is, the total number of photons in all observations is the same regardless of the number of spectral channels M. Typical observations are shown in Fig. 3.



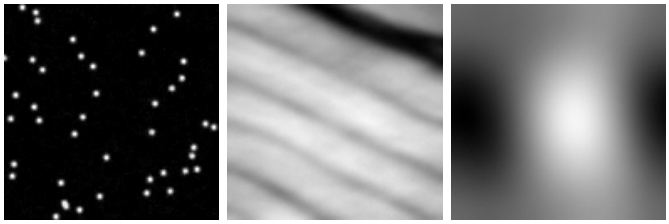
(a) Observation with $M = 20$ (b) Observation with $M = 20$

(c) Observation with $M = 5$ (d) Observation with $M = 5$

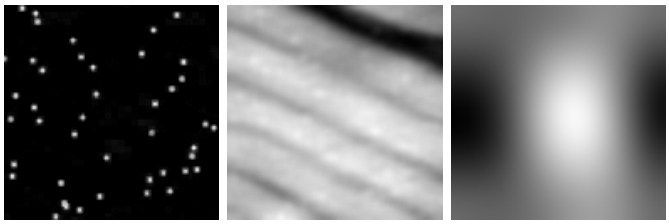
Fig. 3. Typical observations with $T/n = 1000$ using the Case 3 spectra where (a) and (b) are generated using $M = 20$; (c) and (d) are generated using $M = 5$. Bigger M enjoys less coherence but suffers from more noise.

5.2. Simulation Result

We show example reconstructions for two different intensity levels in Fig. 4 ($T/n = 1000$) and Fig. 5 ($T/n = 100$). These show the visual degradation in performance that occurs when the coherence increases. Reconstructions from more observations (lower coherence) show better performance with more accuracy in the muscle component and better shape and brightness in the nuclei component.



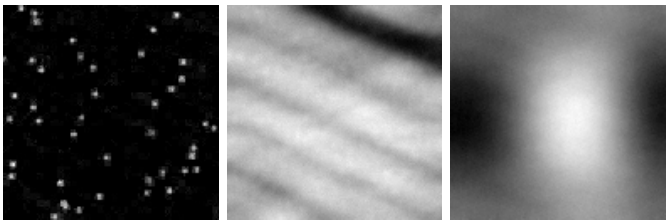
(a) Nuclei Estimate (b) Muscle Estimate (c) Haze Estimate



(d) Nuclei Estimate (e) Muscle Estimate (f) Haze Estimate

Fig. 4. Reconstructions with $T/n = 1000$ where (a)–(c) have low coherence (spectra Case 3 with 20 spectral bands), (d)–(f) have high coherence (spectra Case 3 with 5 spectral bands)

(a) Nuclei Estimate (b) Muscle Estimate (c) Haze Estimate



(d) Nuclei Estimate (e) Muscle Estimate (f) Haze Estimate

Fig. 5. Reconstructions with $T/n = 100$ where (a)–(c) have low coherence (spectra Case 3 with 20 spectral bands), (d)–(f) have high coherence (spectra Case 3 with 5 spectral bands)

In Table 1 shows reconstruction errors with corresponding coherence values for the resulting spectral mixing matrices (A). We use the RMSE expressed as a percentage to evaluate the accuracy of the reconstructions:

$$\text{RMSE}\ (\%) = 100 \cdot \|\widehat{x} - x\|_2 / \|x\|_2 \tag{8}$$

As expected from the theory, with lower coherence (μ) smaller RMSE can be achieved for sufficiently small M. We can also see from the bold entries in this table that the coherence is an excellent predictor for the relative RMSE values.

		Case 1	Case 2	Case 3
$M = 5$	$\mu(A)$	**0.5435**	0.6621	0.9709
RMSE (%)	$T/n = 1000$	**2.4260**	2.5785	5.2896
	$T/n = 100$	**5.9388**	6.3726	8.2410
$M = 12$	$\mu(A)$	0.5077	**0.5426**	0.6108
RMSE (%)	$T/n = 1000$	2.3925	**2.3615**	2.5438
	$T/n = 100$	5.8921	**5.9692**	6.2952
$M = 20$	$\mu(A)$	0.5028	0.5249	**0.5407**
RMSE (%)	$T/n = 1000$	2.3448	2.3466	**2.3307**
	$T/n = 100$	5.8037	5.7696	**5.9125**

Table 1. MSE and coherence results for the experiments. Bold entries highlight settings with similar performance metrics predicted by our theory based on the spectral signatures and the number of spectral bands.

6. CONCLUSIONS

Choosing the ideal number of spectral bands for tissue separation using multispectral observations introduces a myriad of tradeoffs among hardware costs, signal diversity, and photon noise. If M denotes the number of spectral bands observed, then the optimal value of M will depend on both the spectral profiles of the different tissue types which must be distinguished and the total amount of available light, which is proportional to the total data acquisition time. The nature of these tradeoffs can be explored using concepts developed within the compressed sensing literature, and can help optical engineers design cheap yet effective tools for the problem at hand.

7. REFERENCES

[1] N. Chenouard, I. Bloch, and J.-C. Olivo-Marin, "Particle tracking in fluorescent microscopy images improved by morphological source separation," in *Proc. ICIP*, 2009.

[2] M. Raginsky, R. Willett, Z. Harmany, and R. Marcia, "Compressed sensing performance bounds under Poisson noise," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 3990–4002, 2010.

[3] D. Snyder, *Random Point Processes*, Wiley-Interscience, New York, NY, 1975.

[4] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[5] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2191–2202, 2003.

[6] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 15, no. 12, pp. 4203–4215, December 2005.

[7] Z. Harmany, R. Marcia, and R. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms Theory and Practice," *IEEE Transactions on Image Processing*, vol. PP, no. 99, 2011, DOI:10.1109/TIP.2011.2168410.

[8] Z. T. Harmany, J. Mueller, Q. Brown, N. Ramanujam, and R. Willett, "Tissue quantification in photon-limited microendoscopy," in *Proc. SPIE Optics and Photonics*, 2011.